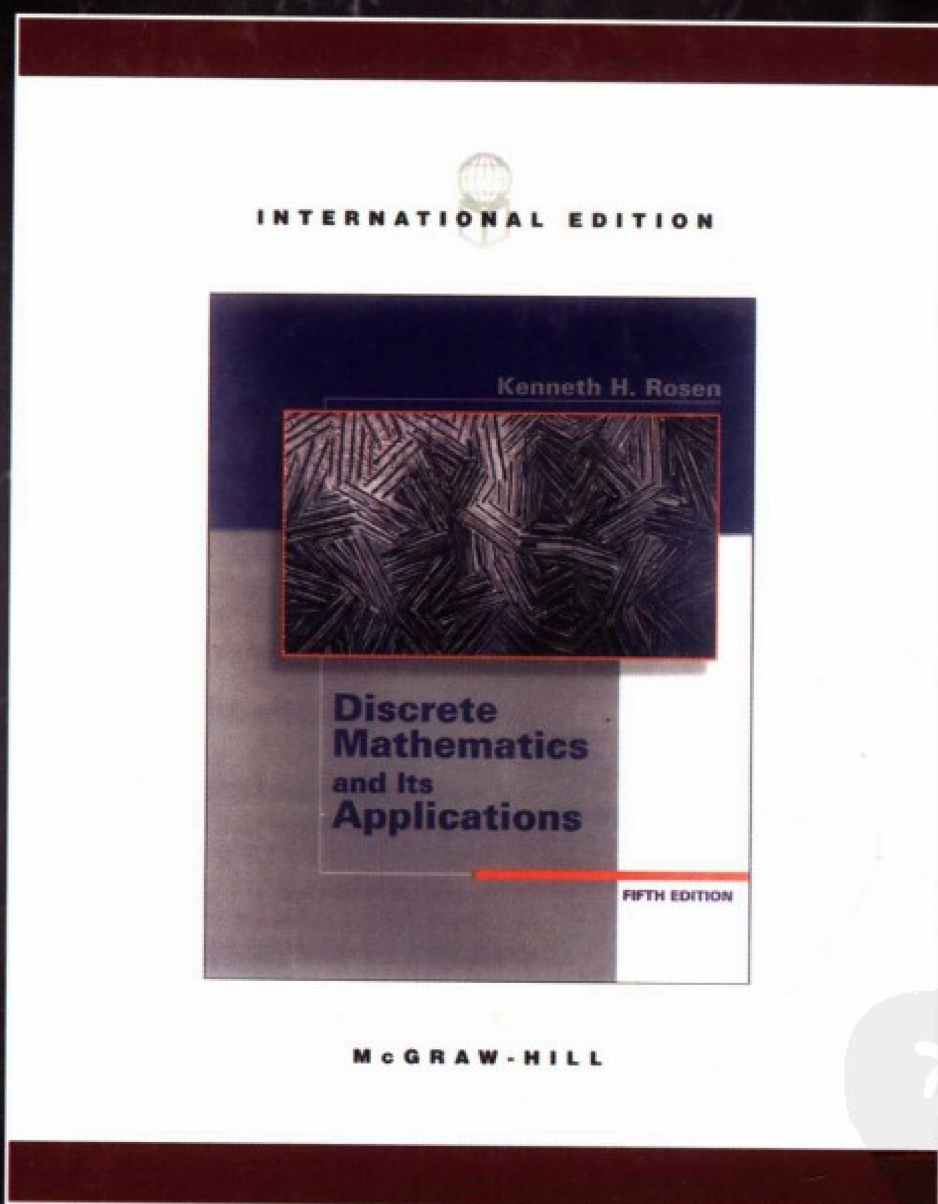
计算机科学丛书

原书第5版

离散数学及其应用

（美）Kenneth H. Rosen 著　袁崇义　屈婉玲　王捍贫　刘田　译



Discrete Mathematics and Its Applications

Fifth Edition



机械工业出版社
China Machine Press

离散数学及其应用（原书第5版）

本书是介绍离散数学理论和方法的经典教材，已经成为全球500多所大学的指定教材，获得了极大的成功。中文版也已被国内大学广泛采用为教材。第5版在前四版的基础上做了大量的改进，使其成为更有效的教学工具。

本书可作为1至2个学期的离散数学课入门教材，适用于数学、计算机科学、计算机工程等专业的学生。

第5版的特点

- 易入门：实践证明本书对初学者来说易读易懂
- 灵活：本教材为灵活使用做了精心设计，各章对其前面内容的依赖降到最小
- 广泛的课堂实践：本书已在500多所大学得到了多年实践检验
- 实例：书中有700多个实例，用于阐明概念，联系不同内容，并引入各种应用
- 应用：本书涉及的应用领域很广，包括计算机科学、数据网络、心理学、化学、工程、语言学、生物学、商业和互联网
- 历史资料：本书以脚注的形式给出了60多位数学家和计算机科学家的传记
- 关键术语和结论：每一章后面都列出了本章的关键术语和结论
- 练习、复习题、补充练习：正文中有3 500多道练习，每章最后都有一组复习题和一组丰富多样的补充练习
- 计算机课题：每一章后面还有一组计算机课题，把学生已经学到的计算和离散数学的内容结合在一起

作者简介

Kenneth H. Rosen 1972年获密歇根大学数学学士学位，1976年获麻省理工学院数学博士学位，1982年加入贝尔实验室，现为AT&T实验室特别成员，国际知名的计算机数学专家，除本书外，还著有《初等数论及其应用》等书。





《离散数学及其应用》（英文版·第5版）
ISBN: 7-111-11503-1
787页，79.00元





上架指导：计算机／离散数学

ISBN 978-7-111-20326-1



投稿热线：(010) 88379604
购书热线：(010) 68995259, 68995264
读者信箱：hzjsj@hzbook.com

封面设计·陈子平

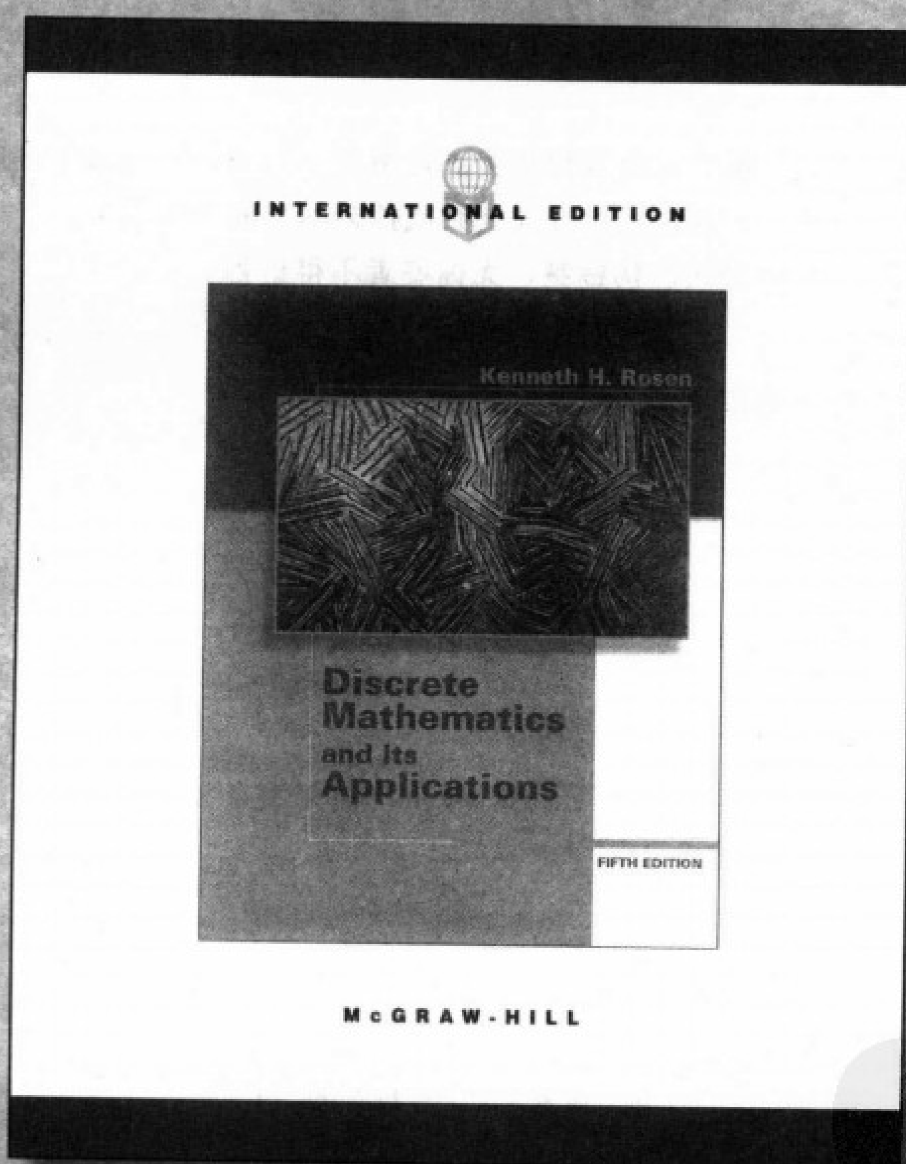
ISBN 978-7-111-20326-1
定价：79.00 元

计 算 机 科 学 丛 书

原书第5版

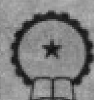
离散数学及其应用

（美）Kenneth H. Rosen 著　袁崇义　屈婉玲　王捍贫　刘田 译



Discrete Mathematics and Its Applications

Fifth Edition



机械工业出版社
China Machine Press

本书是经典的离散数学教材，为全球多所大学广为采用。本书全面而系统地介绍了离散数学的理论和方法，内容涉及数学推理、组合分析、离散结构和算法设计。全书取材广泛，除包括定义、定理的严密陈述外，还配备大量的实例和图表的说明，各种练习和题目，以及丰富的历史资料和网站资源。第5版在前四版的基础上做了大量的改进，使其成为更有效的教学工具。

本书可作为高等院校数学、计算机科学和计算机工程等专业的教材或参考书。

本书版权登记号：图字：01-2003-1003

图书在版编目(CIP)数据

离散数学及其应用(第5版)/(美)罗森(Rosen, K. H.)著；袁崇义等译．-北京：机械工业出版社，2007.6

(计算机科学丛书)

书名原文：Discrete Mathematics and Its Applications, Fifth Edition

ISBN 978-7-111-20326-1

Ⅰ.离…　Ⅱ.①罗…　②袁…　Ⅲ.离散数学-教材　Ⅳ.O158

中国版本图书馆 CIP 数据核字(2006)第133483号

机械工业出版社(北京市西城区百万庄大街22号　邮政编码　100037)

责任编辑：王　玉

三河市明辉印装有限公司印刷·新华书店北京发行所发行

2007年6月第1版第1次印刷

184mm×260mm·51.75印张

定价：79.00元

凡购本书，如有倒页、脱页、缺页，由本社发行部调换

本社购书热线：(010)68326294

出版者的话

文艺复兴以降，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域取得了垄断性的优势；也正是这样的传统，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅擘划了研究的范畴，还揭橥了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短、从业人员较少的现状下，美国等发达国家在其计算机科学发展的几十年间积淀的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章图文信息有限公司较早意识到“出版要为教育服务”。自1998年开始，华章公司就将工作重点放在了遴选、移译国外优秀教材上。经过几年的不懈努力，我们与Prentice Hall，Addison-Wesley，McGraw-Hill，Morgan Kaufmann等世界著名出版公司建立了良好的合作关系，从它们现有的数百种教材中甄选出Tanenbaum，Stroustrup，Kernighan，Jim Gray等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及庋藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力襄助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专程为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍，为进一步推广与发展打下了坚实的基础。

随着学科建设的初步完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都步入一个新的阶段。为此，华章公司将加大引进教材的力度，在“华章教育”的总规划之下出版三个系列的计算机教材：除“计算机科学丛书”之外，对影印版的教材，则单独开辟出“经典原版书库”；同时，引进全美通行的教学辅导书“Schaum's Outlines”系列组成“全美经典学习指导系列”。为了保证这三套丛书的权威性，同时也为了更好地为学校和老师们服务，华章公司聘请了中国科学院、北京大学、清华大学、国防科技大学、复旦大学、上海交通大学、南京大学、浙江大学、中国科技大学、哈尔滨工业大学、西安交通大学、中国人民大学、北京航空航天大学、北京邮电大学、中山大学、解放军理工大学、郑州大学、湖北工学院、中国国家信息安全测评认证中心等国内重点大学和科研机构在计算机的各个领域的著名学者组成“专家指导委员会”，为我们提供选题意见和出版监督。

这三套丛书是响应教育部提出的使用外版教材的号召，为国内高校的计算机及相关专业

的教学度身订造的。其中许多教材均已为M. I. T.，Stanford，U.C. Berkeley，C. M. U. 等世界名牌大学所采用。不仅涵盖了程序设计、数据结构、操作系统、计算机体系结构、数据库、编译原理、软件工程、图形学、通信与网络、离散数学等国内大学计算机专业普遍开设的核心课程，而且各具特色——有的出自语言设计者之手、有的历经三十年而不衰、有的已被全世界的几百所高校采用。在这些圆熟通博的名师大作的指引之下，读者必将在计算机科学的宫殿中由登堂而入室。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑，这些因素使我们的图书有了质量的保证，但我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。教材的出版只是我们的后续服务的起点。华章公司欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方法如下:

电子邮件: hzjsj@hzbook.com
联系电话:（010）68995264
联系地址: 北京市西城区百万庄南街1号
邮政编码: 100037



专家指导委员会

（按姓氏笔画顺序）

尤晋元	王　珊	冯博琴	史忠植	史美林
石教英	吕　建	孙玉芳	吴世忠	吴时霖
张立昂	李伟琴	李师贤	李建中	杨冬青
邵维忠	陆丽娜	陆鑫达	陈向群	周伯生
周克定	周傲英	孟小峰	岳丽华	范　明
郑国梁	施伯乐	钟玉琢	唐世渭	袁崇义
高传善	梅　宏	程　旭	程时端	谢希仁
裘宗燕	戴　葵			



作者介绍

肯尼思 H. 罗森(Kenneth H. Rosen)博士是 AT&T 米德尔顿实验室(新泽西州)技术员工中的杰出成员。

罗森博士于 1972 年获密歇根大学(位于 Ann Arbor)数学学士学位，1976 年获麻省理工学院数学博士学位，其博士论文研究的是数论，导师是哈罗德·斯达克(Harold Stark)。在 1982 年加入贝尔实验室之前，他在科罗拉多大学(位于 Boulder)和俄亥俄州立大学(位于 Columbus)工作过，还在缅因大学(位于 Orono)担任过数学副教授。在 AT&T 实验室工作时，罗森博士还在蒙马斯大学任教，讲授离散数学、编码理论和数据安全的课程。

罗森博士在专业期刊上发表过许多关于数论及数学建模的论文。他的《初等数论及其应用》一书由 Addison-Wesley 公司出版，目前已是第 4 版。本书由 McGraw -Hill 公司出版，也已出第 5 版。这两本教科书均被几百所大学广为采用。他还是《*UNIX*: *The Complete Reference*, *UNIX System V Release 4*: *An Introduction*》和《*Best UNIX Tips Ever*》的作者之一。这两本书都由 Osborne McGraw-Hill 公司出版，销量都达 10 万册以上，并被译成中文、德文、西班牙文和意大利文。罗森博士还是《*Handbook of Discrete and Combinatorial Mathematics*》的编者。该书已于 2000 年由 CRC 出版社出版。罗森博士是 CRC 出版社离散数学丛书的编辑顾问。他还对将数学软件应用到教育领域和专业领域感兴趣，并且已与滑铁卢 MAPLE 软件公司(Waterloo MAPLE Software)合作研究这两个课题。

在贝尔实验室以及 AT&T 实验室期间，罗森博士参与过各种课题的研究，如运筹学研究、计算机及数据通信设备生产线的设计。他还协助规划 AT&T 在多媒体领域未来的产品和服务，包括视频通信、语音识别和合成以及图像网络。他还对 AT&T 使用的新技术作评估。他发明过许多新式服务，已持有或正申请的专利超过 65 项。他的一个有意思的研究项目是辅助评估对 AT&T EPCOT 中心有吸引力的技术。



前　言

本书是根据我多年讲授离散数学的经验和兴趣写成的。对学生而言，我的目的是为他们提供准确而可读的教材，使离散数学的概念和技术得以清晰地介绍。我的目标是向爱怀疑的学生们展示离散数学的实用性，希望为学习计算机科学的学生提供一切必需的数学基础，也希望学数学的学生理解数学概念的重要性，以及这些概念为什么对应用是重要的，还希望本书既能达到这些目标，又不含太多的水分。

对教师（指导者）而言，我的目的是使用成熟的数学教学技术设计一个灵活而全面的教学工具，希望为教师们提供能够高效地、以最适合特定学生特点的方式讲授离散数学的教材。希望本书能够达到这些目标。

我为本教材在过去已经取得的巨大成功而分外高兴。根据成功使用本书的500多所学校的大批师生的反馈和建议，此次第5版进行了许多改进。很多内容有所提高，辅助材料更加丰富，配套网站提供的材料更有帮助性，使师生更容易达到他们的目标。

本教材是为1至2个学期的离散数学入门课程而设计的，适用于数学、计算机科学、工程等许多专业的学生。唯一的先修课要求是大学代数。

离散数学课的目标

离散数学课有多个目标。学生不仅要学会一些特定的数学知识并知道怎样应用，更重要的是，这样一门课应教会学生怎样进行数学思维。为此，本教材特别强调数学推理及用不同的方法解题。本教材有5个重要的主题交织在一起：数学推理、组合分析、离散结构、算法思维、应用与建模。成功的离散数学课程应该努力使这五部分内容相互融合、平衡。

1. *数学推理*：学生必须理解数学推理，以便阅读、理解和构造数学证明。本教材以数理逻辑开篇，在后面证明方法的讨论中，数理逻辑是基础。本书还十分强调数学归纳法，不仅用许多证明的实例进行介绍，还详细地解释了数学归纳法为什么是有效的证明方法。

2. *组合分析*：解题的一项重要技巧是计数或枚举对象。本书中，对枚举的讨论从基本的计数方法着手，重点是用组合分析方法来解决计数问题，而不直接使用公式。

3. *离散结构*：离散数学课应该教会学生如何使用离散结构，即学会如何使用表示离散对象及其之间的关系的抽象数学结构。离散结构包括：集合、置换、关系、图、树和有限状态机等。

4. *算法思维*：有些问题是通过详细说明其算法来求解的。算法在描述后就可构造计算机程序来实现。这一过程中用到的数学部分包括：算法描述、正确性证明以及执行算法所需要的计算机内存和时间的分析。这些内容在本书中均有介绍。算法是用英语[⊖]和一种易于理解的伪码描述的。

5. *应用与建模*：离散数学几乎应用在所有研究领域中。本书既介绍了其在计算机科学和数据网络中的许多应用，也介绍了在其他各种领域中的应用，如化学、植物学、动物学、语言学、地理、商业以及互联网等。这些均是离散数学的实际而重要的应用，而不是编造的。用离散数学来建模是十分重要的问题求解技巧。本书中的一些练习让学生有机会通过自己构造模型来掌握这一技巧。

⊖ 译著中采用汉语。——译者注

第 5 版的修改

本书的第 4 版已在美国的 500 多所学校使用并获得了成功。加拿大的几十所大学以及欧洲、亚洲和大洋洲的一些大学使用第 4 版也获得了成功。虽然第 4 版已经是一本非常有效的教材，但许多教师(包括长期使用者)还是提出了使本书更有效的修改要求。我花了大量的时间和精力来满足这些要求。

修改的结果就是现在的第 5 版，它为教师和学生提供了比第 4 版更多的内容。最重要的是，第 5 版的内容结构已大为改进，从而使本书成为一本更有效的教学工具。有关逻辑、证明方法和证明策略的材料有了实质性的改进，更能帮助学生掌握数学推理。对于学生经常感到困难的内容，增加了解释和例子加以阐述。在习题中还插入了一些新的练习，有通常的练习，也有富于挑战性的练习。还增加了一些与内容密切相关的应用，如在网络及计算机科学中的应用。配套网站在大量开发活动支持下初具规模，现在已经能提供一些工具，学生可用来掌握一些关键的概念和探索离散数学的奇妙世界。

结构的改进

- 第 1 章集中介绍了数学推理，先介绍命题和谓词逻辑，然后是推理规则，最后是基本的证明技术。
- 大 O 及相关记号的讨论直接放在算法复杂性的前面。
- 数列与求和的讨论直接放在数学归纳法的前一节。
- 用单独一节专门介绍二项式系数。
- 用单独一章专门介绍概率论。
- 排序安排在第 2 章，强调更多的排序算法。

逻辑

- 蕴含式的介绍更深入，还介绍了蕴含式的否命题、逆命题和逆否命题。
- 增加一小节介绍逻辑难题。
- 量词分在两节中介绍。
- 将英语或数学命题翻译为逻辑表达式的解释更加详细，反过来的解释亦然。
- 量词否定的介绍更加深入。
- 介绍了求解及 Prolog 中谓词逻辑的使用。
- 描述了逻辑在系统规范中的应用，这是系统工程师、软件工程师和硬件工程师都感兴趣的内容。

书写和理解证明

- 证明方法安排在第 1 章，这样，书中前面的一些章节就可以直接使用这些材料了。
- 明确讨论了唯一性证明。
- 3.1 节进一步探讨了证明策略，扩充了第 1 章中的初步证明策略。
- 用更多的解释和新例子加强了数学归纳法和强数学归纳法的介绍。
- 介绍了结构归纳法。
- 讨论了递归算法的正确性证明。

算法

- 第 2 章介绍了贪心算法。
- 扩充了递归算法的介绍。
- 加强了深度优先和宽度优先算法的介绍。

- 分析或讨论了更多算法的复杂性。
- 也扩充了分治算法及用来分析其复杂性的递推关系式的介绍。
- 增加了快速模指数算法、求最近点对问题算法、赫夫曼(Huffman)编码算法以及找零钱问题的贪心算法。

应用

- 图模型中增加了网页图、电话呼叫图、好莱坞图、熟人关系图和合作图。
- 讨论了网页蜘蛛所用的搜索方法。
- 描述了卡诺图中格雷码的使用。
- 加强了巴克斯范式的讨论。
- 扩展了 n 元关系和关系数据库的讨论。

数论、组合学和概率论

- 更深入讨论了数论中与公钥密码学相关的内容，如伪素数、卡米切尔数及概率素数测试。
- 扩充了有关不同基扩张之间变换的材料。
- 单用一节专门介绍二项式系数和二项式定理。
- 扩充了概率论的内容。许多关键内容都得到了扩充，并都放在同一章介绍。

图和树

- 增加了用归纳法构造 n 方体的解释。
- 哈密顿回路存在性的充分条件的介绍更加详细。
- 讨论了博弈树和最小-最大策略。
- 介绍了赫夫曼编码。
- 扩充了深度优先和宽度优先搜索。

练习

- 增加了 600 多道练习，有普通的，也有富有挑战性的。特别强调有关逻辑和证明以及数学归纳法的新练习。
- 增加了一些练习来平衡偶数编号的练习和带答案的奇数编号的练习。

新的人物传记和历史资料

- 增加了下列人物的传记：亚里士多德(Aristotle)、谢佛(Sheffer)、Smullyan、Rivest、Shamir、Adleman、卡米切尔(Carmichael)、麦卡锡(McCarthy)和赫夫曼(Huffman)。

网站

- 增加了数百个网络资源的链接。
- 在某些关键领域中增加了新的例子。
- 可以访问一些例子和证明的更详细解释。
- 对某些主题，用于自评的工具已经可以使用，主题包括蕴含、量词、证明方法、函数、大 O 记号、数学归纳法和计数问题。
- 在课本中设置图标，指示网站中相应内容的类型。
- 开发了关键算法的交互式演示，便于与课本一起使用。

特点

易入门　实践证明：此课本对初学者来说易读易懂。它的大部分内容只要求学生学过大学代数，不需要其他的预备知识，少数几个涉及微积分的地方也有明确说明。大部分学生应该很容易理解课本中用于表示算法的伪码，不管他们是否学过程序设计语言。本书不要求正规计算机

科学方面的预备知识。

每章都是从易理解的基础开始。基础的数学概念精确地研究之后，其他领域中更难的部分和应用就提出来了。

灵活 课本为灵活使用作了精心设计。各章对其前面内容的依赖都降到最低程度。每一章分成长度大致相等的若干节。每一节又根据内容划分成小节以便教学。教师可以根据这些分块很容易地安排进度。

写作风格 本书的写作风格是直接和实用。使用准确的数学语言，但没有过分的形式化与过分的抽象。在记号和数学命题的词汇间作了精心的平衡。

广泛的课堂实践 本书已在500多所学校使用过，其中400多所使用了不止一次。这些学校师生的反馈意见使第5版比前几版更成功。

数学严密性和准确性 本书中所有定义和定理的陈述都十分详细，所以学生可以欣赏其语言的准确以及数学所需的严密。但证明则是缓慢引入并展开的，每一步都经过了详细论证。本书解释并大量使用了递归定义。

实例 本书使用了700多个例子来阐述概念、建立不同内容之间的关系或导入应用。在大部分例子中，先提出问题，再适当给出它的解。

应用 书中叙述的应用展示了离散数学在解决现实问题中的使用价值，所涉及的范围很广，包括计算机科学、数据网络、心理学、化学、工程、语言学、生物学、商业和互联网。

算法 离散数学的结论常常要用算法来表示，因此，本书每一章都介绍了一些关键算法。这些算法既可以用文字叙述，也可以用更易于理解的结构化伪码来叙述。附录B对伪码作了描述和规范。书中所有算法的计算复杂性也都给出了初步的分析。

历史资料 本书对许多内容的背景作了简要介绍。以脚注的形式给出了60多位数学家和计算机科学家的简短传记。这些传记介绍了对离散数学做出过重要贡献的科学家们的生活、事业及成就。此外，脚注中还包含了大量历史资料，作为正文中历史资料的补充。

关键术语和结果 每一章后面都列出了本章的关键术语和结果，但只列出了学生必须学会的那些最重要的关键术语，而不是在该章中定义的所有术语。

练习 书中包含3500多道练习，提出了大量不同类型的问题。本书不仅提供了足够多简单问题用于练习基本技巧，还提供了大量的中等难度的练习和许多有挑战性的练习。练习的叙述清晰而无歧义，并按难易程度进行了分级。有些特殊的论题还设计成系列练习。这些练习不仅给出了正文中没有介绍的新概念，还使学生可以通过自己的努力发现新思想。

比平均水平稍难的练习用单个星号作标记，相当有挑战性的问题则用两个星来标记。必须用微积分来解的练习也作了明确说明。导出正文要用到的结果的练习则用符号☛识别。课本最后给出了所有奇数号的练习的答案或解题大纲。解题大纲清楚地给出了大多数证明步骤。

复习题 每章最后都有一组复习题。设计这些问题的目的是帮助学生重点学习该章最重要的概念和技术。要回答这些复习题，学生必须写出较长的答案，而只做点计算或简答是不够的。

补充练习集 每章后面都有一组丰富而多变的补充练习。这些练习一般比每节的练习难度大。补充练习强化该章中的概念，并把不同内容更有效地综合起来。

计算机题目 每一章后面还有一组计算机题目。大约有150个这样的题目，把学生可能已经学到的有关计算和离散数学的内容联系起来。如果从数学角度或程序设计角度来看，其难度超过平均水平的计算机题目用一个星标记，特别有挑战性的则用两颗星标记。

计算和研究 每一章的结论部分都有一组计算和研究性问题。要完成这些练习(大约有100道)需要软件工具的帮助，例如学生或教师自己编写的程序，或数学计算软件包如MAPLE或

Mathematica。大部分这样的练习为学生提供了通过计算发现新事实或新思想的机会。(在《用MAPLE研究离散数学及其应用》一书中也讨论了一些这类练习。)

写作题目 每一章后面都有一组应该书面完成的题目。要完成这类题目，学生需要查阅参考数学文献。有些题目在过去的历史上是很重要的，学生需要查找原始资料，其他的题目则是通往新内容和新思想的途径。所有此类练习均向学生展示正文中没有深入探讨的思想。这些题目把数学概念和书面写作的过程结合在一起，以帮助学生探索未来的研究领域。(在《学生解题指南》中可以找到为这些题目准备的参考文献。)

附录 正文有两个附录，附录A介绍指数函数和对数函数，目的是复习在课程中要反复使用的某些基本内容；附录B则介绍正文中用来描述算法的伪码。

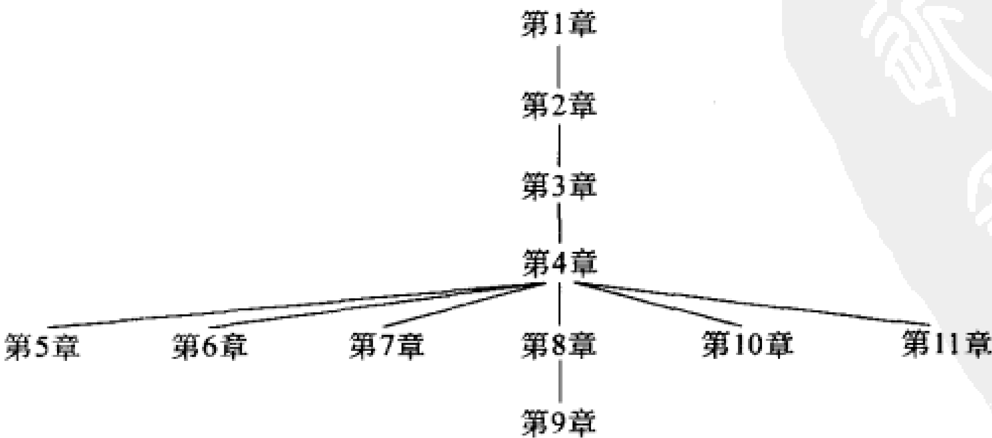
推荐读物 在正文的最后，专门有一节用来为各章推荐参考读物，其中有不超过本书水平的书籍，也有较难的书籍；有综述性文章，也有发表离散数学新发现的原始文章。

怎样使用本书

本书经过了精心写作和编排，适用于不同水平有不同重点的离散数学课。下表列出了核心章节和可选章节。为大学二年级学生开设的一学期的入门课程可以以核心章节为基础，其他章节由教师取舍。两学期的入门课可以在核心章节外加上所有可选的数学章节。强调计算机科学的课程可以使用计算机科学的部分或全部可选章节。

章	核心节	可选的计算机科学章节	可选的数学章节
1	1.1～1.8(根据需要)		
2	2.1～2.4，2.7(根据需要)	2.5	2.6
3	3.1～3.4	3.5，3.6	
4	4.1～4.3	4.6	4.4，4.5
5	5.1	5.3	5.2
6	6.1，6.5	6.3	6.2，6.4，6.6
7	7.1，7.3，7.5	7.2	7.4，7.6
8	8.1～8.5		8.6～8.8
9	9.1	9.2，9.3	9.4，9.5
10		10.1～10.4	
11		11.1～11.5	

使用本书的教师可以选用或略去每节最后有挑战性的例题及练习，以调整课程的难度。每章对以前各章的依赖关系如下图所示。



辅助资料

《学生解题指南》(*Student Solutions Guide*) 这本可以单独购买的学生手册包含了各组练习中所有奇数编号问题的完整解答。这些解答解释了为什么要用某种特定的方法以及为什么这个方法管用。有些问题还给出了一两种其他可能的解法，以说明一个问题可以用不同的方法求解。本指南还包含了为每章后面的写作题目推荐的参考文献，以及对学生书写证明的指导意见，并列出了学生在做离散数学题时常犯的错误。本指南还为每一章提供考试样题及解答，以帮助学生准备考试。学生们感到本指南分外有用。

《教师资料手册》(*Instructor's Resource Guide*) 本手册包含所有偶数编号练习题的完整解答。它对教材的每一章提出了讲解建议，包括每一节中应强调的重点，以及本节内容在整个体系中的位置。此外，本手册为每章提供了考试题库，还包括一些样卷及解答。最后，本手册还给出了教学大纲的样本。

《离散数学应用》(*Applications of Discrete Mathematics*) 这本辅助读物是一本独立的教程，既可结合课本使用，也可独立使用。它包含了使用过课本的教师们撰写的 20 多章内容(每章均有各自的练习)。因为其编排格式与课本类似，所以该书既可作为一门独立课程的课本，也可作为学生讨论班的教材，还可供学生独立研究使用。

试题库 本题库内容广泛，包含了 1600 多道习题，可以在 Windows 系统或 Macintosh 系统中使用。教师可以使用这一软件选择试题，也可随机产生试题。对同样的试题，教师可以加上自己的标题和要求，也可以打乱其顺序后再打印；既可以编排已有的问题，也可以加上自己的问题。在《教师资料手册》中有此题库的打印版，包括试题和解答。

《用 MAPLE 研究离散数学及其应用》(*Exploring Discrete Mathematics and Its Applications with MAPLE*) 这本补充读物用来帮助学生使用计算机代数系统 MAPLE 来完成离散数学的各种计算。对本书的每一章，这本补充读物包括下列内容：相关的 MAPLE 函数及其使用方法的描述，完成相关计算的 MAPLE 程序，每章结尾处还有阐明怎样使用 MAPLE 做计算和研究的建议和例子，以及可用 MAPLE 做的练习题。

配套网站

为本书开发的一个内容广泛的配套网站。该网站将不断地维护和改进。可以多种方式使用该网站进一步学习离散数学，网站地址是：

www. mhhe. com/rosen

由此地址进入该网站的主页。该网页有下列链接：

- 信息中心
- 学生中心
- 教师中心

信息中心：该中心包含本书及其辅助资料的基本信息。教师在这个网页中可以利用网页抽取系统“Page Out”，建立自己的课程网页，还能学习定制出版信息。从该信息中心出发，沿着适当的链接，可以看到本书网站的一个概观。

学生中心：该中心包含了可供学生使用的丰富资源，包括下列与课本紧密相关的资源(在课本中用相应的图标加以标记)：



- 网络资源指南：该指南提供了数百个含相关资料的外部网站的链接。可以通过关

键词浏览或存取这些链接，它们将把你带到包含下列信息的网站：历史及传记信息、谜题及问题、讨论、Java 小程序、程序以及其他类型的资源。



- 额外的例子：这个网站包含了大量额外的例子。这些例子主要集中在学生经常需要的课外资料的领域。虽然大部分例子只是扩充了基本概念，但还有一部分十分具有挑战性。



- 额外的步骤：对一些困难的知识点，提供了更深入的解释帮助理解，尤其是一些特殊的证明和例子。



- 评估：提供对七个关键概念理解程度的评估。每个评估都提供了一个题库，其中的每个试题由两部分构成：先是一段简短的复习，然后是一个多选题。如果你选择的答案不正确，该系统还能提供建议，帮助你理解错在什么地方。这样的评估系统应该能诊断出你学习中的问题，从而把精力集中在寻找纠正办法上。



- 交互式演示：已经开发了八个交互式演示系统，你可以用它们考查一些重要算法是怎么工作的。这些演示都与课本中的相应材料相对应。

从该学生中心你可以访问网络教学系统“Net Tutor”。它提供了在线教学帮助。当你提问与课本内容相关的问题时，如果是在规定教学时间内，你会收到实时回答；否则稍后才会收到回答。

学生中心还支持能发布消息的公告板系统。使用该系统，你可以提出问题，还可以回答其他学生提出的问题。

除此之外，学生中心还包括下列资源：

- 证明书写指南
- 离散数学的常见错误
- 写作题目的建议
- MAPLE 软件

教师中心：除了包含学生中心和信息中心提供的所有链接外，网站中的教师中心还包含下列链接：

- 教学大纲样本
- 教学建议
- 《*Applications of Discrete Mathematics*》一书中的某些章节

写给学生

什么是离散数学？离散数学是数学中研究离散对象的部分。（这里“离散”的含义是“由不同的或不相连的元素组成”。）离散数学能解决的问题包括：

- 在计算机系统中，有多少种方式可以选择一个合法口令？
- 赢彩票的概率是多少？
- 两台计算机之间在网络上是否有通路？
- 在某一交通系统下，两个城市之间的最短路径是什么？
- 怎样把整数序列按递增序排列？

- 完成上述排序需要多少步骤？
- 如何证明一个排序方法能正确地排序？
- 怎样设计两个整数相加的电路？
- 有多少合法的因特网网址？

你们将学习解决诸如以上问题要用到的离散结构和技术。

更一般地，在对对象进行计数时要用到离散数学，研究两个有限(或可数)集合之间的关系时要用到离散数学，分析只含有限步的进程时也要用到离散数学。离散数学的重要性还在不断增长，一个关键原因就是计算机以离散的方式存储和处理信息。

为什么要学离散数学？ 有几条重要的理由需要学习离散数学。首先，通过这个课程你们可以发展自己的数学素质，即理解和创造数学证明的能力。没有这些技巧，你们在学习数学时不可能太深入。

第二，离散数学是学习所有更高级数学课程的必经之路。离散数学为学习计算机科学课程提供必要的数学基础，这些课程如：数据结构、算法、数据库理论、自动机理论、形式语言、编译理论、计算机安全以及操作系统。学生如果没有适当的离散数学基础，在学习上述课程时会感到很困难。有个学生给我发电子邮件说，在她学习的每一门计算机科学的课中都用到了本书的知识。

以离散数学为基础的数学课程包括逻辑、集合论、数论、线性代数、抽象代数、组合论、图论及概率论(其离散部分)。

此外，离散数学还包括了用来解运筹学问题(包括许多离散优化技术)、化学问题、工程问题及生物学问题等必要的数学背景。从本书中你将学到在上述某些领域中的应用。

许多学生都感到，与他们以前学过的课程相比，离散数学入门课程的挑战性大得多。这是因为，本课程的一个主要目的是教你进行数学推理和问题求解，而不只是一些分散的技巧。从课本中练习的设计可以看出这个目的。课本中虽然有大量与重点例子类似的练习，但还是有相当比例的练习需要创造性思想。这是有意设计的。虽然课本中的材料提供了解这些问题的工具，但你的任务是创造性地使用这些工具并取得成功。本课程的另一个主要目的是学会处理你以前没有见过的问题。然而，只学会解特殊类型的练习还无法保证能学会足够多的解题技巧，也不能保证在后继课程的学习中或在将来的职业生涯中取得成功。虽然课本讨论了许多主题，但离散数学是一个极为广泛且充满变化的研究领域。作为作者，我的任务之一是帮助你开发学习新知识的能力，在将来的奋斗中你十分需要新的知识。

练习 我愿意就如何学好离散数学(或数学的其他学科和计算机科学)给同学们提点有益的建议。积极地做练习能使你最大地获益。我建议你尽可能地多做练习。在做完老师布置的练习后，我鼓励你做更多的练习，包括本书每节后面的练习和每章后面的补充练习。(注意练习前面的分级标记。)

练习标记	含　　义
无标记	常规练习
*	较难的练习
**	富有挑战性的练习
☞	正文中要用到该练习的某个结论
(需要用到微积分)	解题时要用到极限或微积分的概念

解题时，最好在查阅书后的答案或《学生解题指南》中的答案以前，自己先进行尝试。书后提

供了所有奇数编号练习的答案。注意，只是答案，而不是完整解答。特别地，答案中省略了解的推导过程。《学生解题指南》包括了课本中所有奇数编号练习的完整解答。在解奇数编号练习时，只有当你身陷绝境时，才建议查阅《学生解题指南》寻找解题指导。你尝试得越多，而不只是被动查阅或抄袭解答，你学到的就越多。出版商有意不提供偶数编号的练习的答案和解答。在解这些练习时，若遇到问题就问老师。

网络资源 我强烈推荐你利用网络提供的新资源，特别是专为本书设计的网站 http://www.mhhe.com/rosen 中的资源。其中有许多额外的例子，有为学生经常会遇到的问题而补充的额外步骤，有测量你对关键概念理解程度的评估，有考察关键算法的生动演示，还有一个指向外部网站的链接的大杂烩，通过它们可以探索离散数学的奇妙世界。还有一个能和其他学生进行讨论的公告板系统，可以用它请求其他学生帮助你理解概念或指导解题。帮助其他人的学生会发现，回答问题的同时也帮助了自己更好地掌握内容。你甚至还能访问到一个在线教学网站，可以和老师进行实时交流，得到老师的帮助。（关于使用他人帮助结果的问题，请征求老师的意见。）有关这个网站的详细信息，请参阅前面“配套网站”的描述。

本书的价格 最后，我了解到本书的价格偏高，但我希望你在本书的投资能得到优质的回报。我们花了多年的努力来开发和优化本书及其配套的辅助读物和网站。对大多数人来说，我很自信本书及其配套资料对掌握离散数学大有帮助。即使你现在的课程可能略去了其中某些章节，但当你学习新课程时就会发现，阅读本书中的相关章节对新课程仍然十分有益，许多学生都有这样的感觉。许多人会在将来的研究工作时又找到本书，把它当做一本有用的工具书，特别是那些继续学习计算机科学、数学或工程的人。

致谢

我要感谢很多学校中使用本书的大批师生们，是他们向我提出了宝贵的反馈和有益的建议。没有他们的反馈和建议，本书不可能有很大的改进。我要特别感谢 Jerrold Grossman 和 John Michaels，感谢他们对第 5 版的技术审阅，他们敏锐的目光确保了本书的准确性。我也要感谢通过网站提交评论的人们提供的帮助。

我感谢本书第 5 版及前面四版的许多评阅人，他们对我提出了许多有益的批评和鼓励，我希望本版将不辜负他们的期望。

第 5 版评阅人名单

Kendall Atkinson,
University of Iowa, Iowa City

Zhaojun Bai,
University of California, Davis

Klaus Bichteler,
University of Texas, Austin

Scott Buffett,
University of New Brunswick

E. Rodney Canfield,
University of Georgia

George J. Davis,
Georgia State University

Cary Lee,
Grossmont Community College

Bruce S. Elenbogen,
University of Michigan, Dearborn

Jonathan Goldstine,
Pennsylvania State University

Brian Gray,
Howard Community College

Jonathan Gross,
Columbia University

Jerrold Grossman,
Oakland University

David F. Hayes,
San Jose State University

Truc T. Nguyen,
Bowling Green State University

Akihiro Kanamori,
Boston University

Takashi Kimura,
Boston University

Shui F. Lam,
California State University, Long Beach

Harbir Lamba,
George Mason University

Sheau Dong Lang,
University of Central Florida

Robert Rodman,
North Carolina State University

Stephen C. Locke,
Florida Atlantic University

George F. Luger,
University of New Mexico

John G. Michaels,
SUNY Brockport

Thomas D. Morley,
Georgia Tech

Timothy S. Norfolk, Sr.
University of Akron

George Novacky,
University of Pittsburgh

Jaroslav Opatrny,
Concordia University

Jonathan Pakianathan,
University of Rochester

Halina Przymusinska,
California State Polytechnic University, Pomona

Don Reichman,
Mercer County Community College

Matthew J. Saltzman,
Clemson University

Michael J. Schlosser,
The Ohio State University

Alistair Sinclair,
University of California, Berkeley

Carl H. Smith,
University of Maryland

Patrick Tantalo,
University of California, Santa Cruz

前 4 版评阅人名单

Eric Allender,
Rutgers University

Stephen Andrilli,
La Salle University

Jack R. Barone,
Baruch College

Alfred E. Borm,
Southwest Texas State University

Ken W. Bosworth,
University of Maryland

Lois Brady,
California Polytechnic State University, San Luis Obispo

Russell Campbell,
University of Northern Iowa

Kevin Carolan,
Marist College

Tim Carroll,
Bloomsburg University

Kit C. Chan,
Bowling Green State University

Allan C. Cochran,
University of Arkansas

Peter Collinge,
Monroe Community College

Ron Davis,
Millersville University

Nachum Dershowitz,
University of Illinois, Urbana–Champaign

Charles Parry,
Virginia Polytechnic Institute and State University

Thomas Dowling,
The Ohio State University

Patrick C. Fischer,
Vanderbilt University

Jane Fritz,
University of New Brunswick

Ladnor Geissinger,
University of North Carolina

Paul Gormley,
Villanova University

Laxmi N. Gupta,
Rochester Institute of Technology

Daniel Gusfield,
University of California at Davis

Xin He,
SUNY at Buffalo

Arthur M. Hobbs,
Texas A&M University

Donald Hutchison,
Clackamas Community College

Kenneth Johnson,
North Dakota State University

David Jonah,
Wayne State University

W. Thomas Kiley,
George Mason University

Nancy Kinnersley,
University of Kansas

Alyssa Sankey,
Slippery Rock University

Dinesh Sarvate,

Gary Klatt,
University of Wisconsin

Nicholas Krier,
Colorado State University

Lawrence S. Kroll,
San Francisco State University

Robert Lavelle,
Iona College

Yi-Hsin Liu,
University of Nebraska at Omaha

George Luger,
University of New Mexico

David S. McAllister,
North Carolina State University

Robert McGuigan,
Westfield State College

Michael Maller,
Queens College

Ernie Manes,
University of Massachusetts

Francis Masat,
Glassboro State College

J. M. Metzger,
University of North Dakota

D. R. Morrison,
University of New Mexico

Ho Kuen Ng,
San Jose State University

Jeffrey Nunemacher,
Ohio Wesleyan University

Philip D. Tiu,
Oglethorpe University

Lisa Townsley-Kulich,

Thomas W. Parsons,
Hofstra University
Mary K. Prisco,
University of Wisconsin–Green Bay
Harold Reiter,
University of North Carolina
Adrian Riskin,
Northern Arizona State University
Amy L. Rocha,
San Jose State University
Janet Roll,
University of Findlay
College of Charleston
Steven R. Seidel,
Michigan Technological University
Douglas Shier,
Clemson University
Hunter S. Snevily,
University of Idaho
Daniel Somerville,
University of Massachusetts, Boston
Bharti Temkin,
Texas Tech University
Wallace Terwilligen,
Bowling Green State University
Illinois Benedictine College
George Trapp,
West Virginia University
David S. Tucker,
Midwestern State University
Thomas Upson,
Rochester Institute of Technology
Roman Voronka,
New Jersey Institute of Technology
James Walker,
University of South Carolina
Anthony S. Wojcik,
Michigan State University

我还要感谢 McGraw-Hill 出版社高等教育组的工作人员对本课题强有力的支持。特别要感谢出版人 Bill Barter 的支持，感谢执行编辑 Robert Ross 的提倡和热情，感谢开发编辑 Michelle Munn 的献身和关注，还要感谢原编辑 Wayne Yuhasz，是他的洞察力和技巧保证了本书的成功。同样要感谢本版的策划编辑 Maggie Rogers，是他协助策划并启动了本版的工作。我和他已经共事了很多年。

我还要对制作本书第 5 版做出过贡献的职员表示感谢，他们有：高级项目经理 Joyce Berendes、设计师 K. Wayne Harms、网站开发者 Jeff Huettman、补充读物协调人 Laura Fuller 以及市场经理 Mary Kittell。我还很感谢 Georgia Kamvosoulis Mederer，他检查了全部初稿以保证准确性，感谢校对员 Mary Ellen Oliver，感谢 Paul Lorczak，他检查了题解手册部分及习题答案部分的准确性。

我一直十分感激我在 AT&T 实验室的领导对我的支持。他们为我的专业发展提供了必要的环境，并无私地给我提供必需的资源，没有这些支持也就没有本书的成功。

Kenneth H. Rosen



目　　录



第 1 章　基础：逻辑和证明、集合、函数

本章介绍离散数学的基础知识。有三个主要内容：逻辑、集合和函数。逻辑规则给出数学语句的准确含义。例如，逻辑规则有助于我们理解下列语句及其推理："存在一个不是两个整数的平方和的整数"，以及"对每个整数 n，小于等于 n 的正整数之和是 $n(n+1)/2$"。逻辑是所有数学推理的基础，对计算机的设计、系统规范说明、人工智能、计算机程序设计、程序设计语言以及计算机科学的其他领域，逻辑都有实际的应用。

为了理解数学，就必须理解正确的数学论证(即证明)是由什么组成的。为了学习数学，就需要构造数学论证而不仅仅是阅读说明。本章的目的在于讲授正确的数学论证是由什么组成的，并且介绍构造这些论证的工具。证明不仅对于数学是重要的，而且在计算机科学的很多方面也是重要的，包括程序验证、算法正确性和系统安全性等。而且，自动推理系统已经被构造出来，允许计算机构造自己的证明。

离散数学的不少内容是研究用于表示离散对象的离散结构的。许多重要的离散结构都是用集合构造的，而集合则指的是一组对象。从集合构造的离散结构例子包括：计数时广为应用的对象组合，也就是无序的对象集；表示对象之间相互关连的关系，也就是有序偶的集合；图形，也就是顶点集合及连接顶点的边的集合；以及用于模拟计算机的有限状态机。

函数概念是离散数学中特别重要的概念。函数为一个集合中的每个元素恰好指派另一个集合中的某个元素。像序列和字符串等这些有用的结构都是函数的特例。函数在整个离散数学中扮演着重要的角色，主要用于表示算法的计算复杂性，研究集合的大小，对不同类别的对象计数，以及不计其数的其他方面。

1.1　逻辑

1.1.1　引言

逻辑规则给出数学语句的准确含义，这些规则用来区分有效和无效的数学论证。由于本书的一个主要目的是教会读者如何理解和如何构造正确的数学论证，所以我们从介绍逻辑开始离散数学的学习。

逻辑不仅对理解数学推理十分重要，而且在计算机科学中有许多应用。这些逻辑规则用于计算机电路设计、计算机程序构造、程序正确性证明以及许多其他方面。在随后的几章中将逐一讨论这些应用。

1.1.2　命题

我们首先介绍逻辑的基本成分——命题。*命题*是一个或真或假的陈述语句，但不能既真又假。

例 1　下面的陈述句均为命题。

1. 华盛顿是美国的首都。
2. 多伦多是加拿大的首都。
3. $1+1=2$。
4. $2+2=3$。

命题 1 和 3 成真，命题 2 和 4 为假。■



下一例子给出了不是命题的若干语句。

例 2 考虑下述语句：

1. 几点了？
2. 仔细读这个。
3. $x+1=2$。
4. $x+y=z$。

语句 1 和 2 不是命题，因为它们不是陈述语句。语句 3 和 4 不是命题，因为它们既不成真，也不为假，这是由于语句中的变量没有被赋值。在 1.3 节将讨论这一类语句形成命题的多种方法。 ■

命题常用字母来表示，就像用字母表示变量那样。习惯上用来表示命题的字母是 p，q，r，s，…等。如果一个命题是真命题，它的*真值*为真，用 T 表示；如果它是假命题，其真值为假，用 F 表示。

涉及命题的逻辑领域称为*命题演算*或*命题逻辑*。它最初是 2300 年前由古希腊哲学家亚里士多德[⊖]系统地创建的。

现在我们转而注意从那些已有的命题产生新命题的方法，这些方法在 1854 年曾由英国数学家布尔[⊖]在他的题为《*The Laws of Thought*》的书中讨论过。许多数学陈述都是由一个或多个命题组合而来。称为*复合命题*的新命题由已有的命题用逻辑运算符组合而来。

定义 1 令 p 为一命题，则语句

“不是 p 所说的情形。”

是另一个命题，称为 p 的*否定*。p 的否定用 $\neg p$ 表示。命题 $\neg p$ 读作“非 p”。

例 3 找出命题

“今天是星期五。”

的否定，并用中文表示。

⊖ **亚里士多德(公元前 384—公元前 322)**生于希腊北部的斯塔基尔地区。他的父亲是马其顿国王的宫廷侍医。因为父亲早年便去世了，亚里士多德没能子承父业。当他的母亲也去世后，年轻的亚里士多德就成了孤儿。他的监护人抚养他长大，并教授他诗歌、修辞艺术和希腊语。在亚里士多德 17 岁时，监护人将他送到雅典进一步深造。此后 20 年里亚里士多德在雅典柏拉图学院跟柏拉图学习并进行自己的学术研究。当柏拉图于公元前 347 年去世时，亚里士多德没有选择继承师钵，因为他的哲学思想与柏拉图有很大的分歧。此后，亚里士多德进入 Hermeas 国王的宫廷供职达三年，并与国王的侄女结了婚。当波斯人打败 Hermeas 国王后，他逃往 Mytilene 并受马其顿国王腓力二世的聘请，担任起太子亚历山大(就是后来著名的亚历山大大帝)的老师。在腓力二世逝世后，亚里士多德重返雅典并成立了“吕克昂”学园。这个学派的老师和学生们习惯在花园中边散步边讨论问题，因而得名为“逍遥派”。亚里士多德在“吕克昂”学园讲学达 13 年，他早上给学园里的优秀学生讲课，晚上则给广大听众演讲。当亚历山大大帝于公元前 323 年去世后，那里立刻掀起了反马其顿的狂潮，雅典人攻击亚里士多德，并判他为不敬神罪，亚里士多德最终逃出了雅典，但第二年他就因病去世了，终年 63 岁。

亚里士多德的著作主要分为三类：一类是公开发表的供一般人阅读的文集，一类是实践科学汇编，一类是系统的论辩文集。论辩文集涉及逻辑学、哲学、心理学、医学和自然历史。

⊖ **乔治·布尔(George Boole，1815—1864)**是皮匠的儿子，1815 年 11 月生于英格兰的林肯。由于家境贫寒，布尔不得不在协助养家的同时为自己能受教育而奋斗。不管怎么说，他成了 19 世纪最重要的数学家之一。尽管他考虑过以牧师为业，但最终还是决定从教，并且不久就开办了自己的学校。在备课的时候，布尔不满意当时的数学课本，便决定阅读伟大数学家的论文。在阅读伟大的法国数学家拉格朗日的论文时，布尔有了变分方面的新发现。变分是数学分析的一个分支，它处理的是寻求优化某些参数的曲线和曲面。

1848 年，布尔出版了《*The Mathematical Analysis of Logic*》，这是他对符号逻辑诸多贡献中的第一次。1849 年，他被任命为位于爱尔兰科克的皇后学院的数学教授。1854 年，他出版了《*The Laws of Thought*》，这是他最著名的著作。在这本书中布尔介绍了现在以他的名字命名的布尔代数。布尔撰写了微分方程和差分方程的课本，这些课本在英国一直使用到 19 世纪末。布尔在 1855 年结婚，他的妻子是皇后学院一位希腊文教授的侄女。1864 年，布尔死于肺炎，肺炎是他在暴风雨天气中尽管已经湿淋淋了，但仍坚持上课引起的。

Extra Examples **解** 否定为

“并非今天是星期五。”

也可以更简单地表达为

“今天不是星期五。” ■

注意 严格地说，像例3这种含有可变时间的语句不是命题，除非假定了一个确定的时间。同样，除非假定了确定的地点，否则含有可变地点的语句不是命题；除非假定了确定的人，否则含有可变代词的语句不是命题。

真值表给出命题真值之间的关系。在确定由较简单命题组成的命题之真值时，真值表特别有用。表1-1给出的是命题p及其否定$\neg p$所有可能的真值。

表 1-1 命题之否定的真值

p	$\neg p$
T	F
F	T

命题的否定也可以看作非运算符作用在命题上的结果。非运算符从一个已有的命题构造出一个新命题。现在将引入从两个或多个已有命题构造新命题的逻辑运算符，这些逻辑运算符也称为联接词。

定义 2 令p和q为命题。用$p \wedge q$表示的命题“p而且q”是这样一个命题：当p和q均成真时它成真，否则为假。命题$p \wedge q$称为p和q的合取。

表1-2给出了$p \wedge q$的真值表。注意真值表中共有四行，每行对应着命题p和q真值的一种组合。

例 4 找出命题p和q的合取，其中p为命题“今天是星期五”，q为命题“今天下雨”。

解 这两个命题的合取$p \wedge q$是命题“今天是星期五而且下雨”。这一命题在下雨的星期五成真，不是星期五的日子为假，不下雨的星期五也为假。 ■

定义 3 令p和q为命题，用$p \vee q$表示的命题“p或q”是这样一个命题：它的真值在p和q均为假时为假，否则成真。命题$p \vee q$称为p和q的析取。

$p \vee q$的真值表如表1-3所示。联接词“或”在析取中的使用对应于词“或”(or)包含的两种情况之一，即是“同或”(inclusive or)。析取所含两命题之一成真或两者均成真时，析取的真值为真。例如，下面这句话中，“或”表达的即是“同或”：

“选修过微积分或计算机科学的学生可以选修本课。”

这里我们指的是，选修过微积分和计算机科学两门课的学生以及只选修过其中一门课的学生都可以选修本课。另一方面，当我们说：

“学过微积分或学过计算机科学，但不是两者都学过的学生，可以注册本课。”的时候，使用的“或”是“异或”。这里我们的意思是既学过微积分，又学过计算机科学的学生不能选修本课；只有那些恰好在这两门课中选修过一门的学生可以选修本课。

表 1-2 两命题合取的真值表

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

表 1-3 两命题析取的真值表

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

同样，若餐馆的菜单上写着“汤或沙拉，加一道小菜。”一般这都表示顾客可以选择汤，也可以选择沙拉，但不是既选汤又选沙拉时，才加一道小菜。因此，这里是“异或”而不是“同或”。

例 5 若 p 和 q 就是例 4 中的两个命题，它们的析取是什么？

解 p 和 q 的析取 $p \vee q$ 是命题

“今天是星期五，或今天下雨。”

这一命题在星期五或下雨天(包括下雨的星期五)的任何一天都成真。只有在既不是星期五，又不下雨的日子，此命题为假。 ■

Extra Examples 前面已经说过，在析取中使用的联接词“或”(or)对应于两种情形之一，即“同或”。所以当析取中的两个命题之一成真或两者均成真时，析取成真。有时我们也按“异或”的含义使用“或”。用异或来联结命题 p 和 q 时，就得到命题“p 或 q(但非两者)”。这一命题当 p 成真且 q 为假时成真，或反过来当 p 为假且 q 成真时也成真。p 和 q 两者均为假或均成真时，这一命题为假。

定义 4 令 p 和 q 为命题。p 和 q 的异或，用 $p \oplus q$ 表示，是这样一个命题：当 p 和 q 中恰有一个成真时它成真，否则它为假。

两个命题异或的真值表如表 1-4 所示。

1.1.3 蕴含

我们还将讨论其他几个重要的命题组合方式。

Assessment **定义 5** 令 p 和 q 为命题。蕴含 $p \rightarrow q$ 是这样一个命题：在 p 成真而 q 为假时它为假，否则成真。在这一蕴含命题中，p 称为假设(或前项，前提)，q 称为结论(或推论)。

蕴含命题 $p \rightarrow q$ 的真值表如表 1-5 所示。一个蕴含有时也称为条件语句。

表 1-4 两命题异或的真值表

p	q	$p \oplus q$
T	T	F
T	F	T
F	T	T
F	F	F

表 1-5 蕴含 $p \rightarrow q$ 的真值表

p	q	$p \rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

由于数学推理中许多地方出现蕴含，表示 $p \rightarrow q$ 的术语很多，下面是常用的几个：

Extra Examples

- “如果 p，那么 q”
- “如果 p，则 q”
- “p 仅当 q”
- “p 成立，则 q 成立”
- “q 的充分条件是 p”
- “q 由 p 推出”
- “p 是 q 的充分条件”
- “p 蕴含 q”
- “q 如果 p”
- “q 每当 p”
- “p 的必要条件是 q”
- “q 是 p 的必要条件”

注意，只有在 p 成真而 q 为假时 $p \rightarrow q$ 才为假，所以当 p 和 q 均成真，或 p 为假(无论 q 的真值是什么)时，其真值为真。

有助于记住“当蕴含命题的前提为假时其值为真”这一事实的方法是把蕴含想像为合同或义务。例如，许多政治家在竞选时都许诺：

“如果我当选了，那么我将会减税。”

如果这个政治家当选了，选民将期望他能减税。进一步地，如果这个政治家没有当选，那么选民就无法指望他能减税，尽管这个人也许有足够的影响力可令当权者减税。只有在该政治家当选但却没有减税的情况下，选民才能说政治家违背了竞选诺言。这种情形对应于在 $p \rightarrow q$ 中 p 为真，但 q 为假的情况。

类似地，考虑教授可能作出的如下陈述：

“如果你在期末考试得了满分，那么你的成绩将评定为A。”

如果你设法在期末考试得了满分，那么你可以期望得到A。如果你没得到满分，那么你是否能得到A将取决于其他因素。然而，如果你得到满分，但教授没有给你A，你会有受骗的感觉。

许多人对“p 仅当 q”与“如果 p，那么 q”表示同一意思感到有些不解。要理解这一点，请记住，“p 仅当 q”说的是当 q 不为真时 p 不能为真。也就是说，如果 p 为真但 q 为假，则这个语句为假。当 p 为假时，q 或者为真，或者为假，因为语句并没有谈及 q 的真值。人们常犯的一个错误是用“q 仅当 p”来表达 $p \rightarrow q$。然而，当 p 和 q 有不同的真值时，这两个语句有不同的真值。

我们定义蕴含的方式比语言中的蕴含含义更广泛一些。例如，在

“如果今日天晴，那么我们将去海滩。”

中有假设和结论之间的联系，这是一般语言中的蕴含。而且除非今日的确天晴但我们不去海滩，否则上述蕴含总是成立。另一方面，根据蕴含命题的定义，蕴含语句

“如果今天是星期五，那么 $2+3=5$。”

总是成立的，因为它的结论是成真的（于是假设部分的真值不起作用）。蕴含语句

“如果今天是星期五，那么 $2+3=6$。”

是除星期五以外天天成真，尽管 $2+3=6$ 为假。

在自然语言中，我们不会使用最后这两个蕴含命题（除非偶尔在讽刺中用到），因为其中的假设和结论之间没有什么联系。在数学推理中我们考虑的蕴含命题比语言中使用的要广泛一些。蕴含作为一个数学概念不依赖于假设和结论之间的因果关系。我们关于蕴含的定义规定了它的真值，而这一定义不是以语言的用法为基础的。

许多程序设计语言中使用的 if-then（如果-那么）结构与逻辑中使用的不同。大部分程序设计语言中都有 **if** p **then** S 这样的语句，其中 p 是命题而 S 是一个程序段（待执行的一条或多条语句）。当程序的运行遇到这样一条语句时，如果 p 为真，就执行 S；但若 p 为假，则 S 不执行。下面的例子说明这一点。

例 6 若执行语句

if $2+2=4$ **then** $x:=x+1$

之前，$x=0$，执行以后 x 的值是什么？（符号 $:=$ 代表赋值，语句 $x:=x+1$ 表示将 $x+1$ 的值赋给 x。）

解 因为 $2+2=4$ 为真，赋值语句 $x:=x+1$ 被执行。因此在执行此语句之后 x 的值是 $0+1=1$。 ■

逆、倒置与反

由 $p \rightarrow q$ 可以构成几个相关的蕴含。命题 $q \rightarrow p$ 称为 $p \rightarrow q$ 的**逆蕴含**，而 $p \rightarrow q$ 的**倒置蕴含**是命题 $\neg q \rightarrow \neg p$。命题 $\neg p \rightarrow \neg q$ 称为 $p \rightarrow q$ 的**反蕴含**。

蕴含 $p \rightarrow q$ 的倒置 $\neg q \rightarrow \neg p$ 与 $p \rightarrow q$ 有相同的真值。要明白这一点，注意仅当 $\neg p$ 为假且 $\neg q$ 为

真时也就是仅当 p 为真且 q 为假时，倒置为假。另一方面，逆蕴含 $q\rightarrow p$ 和反蕴含 $\neg p\rightarrow\neg q$ 都不与 $p\rightarrow q$ 具有相同的真值。要明白这一点，注意当 p 为真且 q 为假时，原来的蕴含 $p\rightarrow q$ 为假，但逆蕴含与反蕴含都为真。当两个复合命题总是具有相同的真值时，我们称它们等价。因此一个蕴含与其倒置等价。一个蕴含的逆蕴含与反蕴含也是等价的，读者能够验证这一点。(1.2 节将介绍等价命题。)一个常见的逻辑错误是假设一个蕴含的逆蕴含和反蕴含等价于这个蕴含。

例 7 说明蕴含的用法。

Extra Examples **例 7** 找出如下蕴含的倒置蕴含、逆蕴含和反蕴含：

"每当下雨时，主队就能获胜。"

解 因为"q 每当 p"是表达蕴含 $p\rightarrow q$ 的一种方式，原始语句可以被改写为：

"如果下雨，那么主队就能获胜。"

因此，这个蕴含的倒置蕴含是

"如果主队没有获胜，那么没有下雨。"

逆蕴含是

"如果主队获胜，那么下雨了。"

反蕴含是

"如果没有下雨，那么主队没有获胜。"

其中只有倒置蕴含等价于原始语句。 ■

定义 6 令 p 和 q 为命题。*双蕴含* $p\leftrightarrow q$ 是这样一个命题：其真值只有在 p 和 q 有同样的真值时为真，否则为假。

$p\leftrightarrow q$ 的真值表如表 1-6 所示。注意，双蕴含 $p\leftrightarrow q$ 恰在 $p\rightarrow q$ 和 $q\rightarrow p$ 两个蕴含均为真时为真。正因为如此，术语

"p 当且仅当 q"

常用来表示这一双蕴含，并且在符号书写上是把符号 → 和 ← 结合起来。表达命题 $p\leftrightarrow q$ 的其他方式是："p 是 q 的充分必要条件"，"如果 p，那么 q；反之亦然"以及"p 如果 q"。表示双蕴含的最后一种方式是用缩写符号"iff"代替"当且仅当"。注意，$p\leftrightarrow q$ 与 $(p\rightarrow q)\wedge(q\rightarrow p)$ 有完全相同的真值。

表 1-6 双蕴含 $p\leftrightarrow q$ 的真值表

p	q	$p\leftrightarrow q$
T	T	T
T	F	F
F	T	F
F	F	T

Extra Examples **例 8** 令 p 为语句"你可以坐飞机"，令 q 为语句"你买机票了"。则 $p\leftrightarrow q$ 为语句

"你可以坐飞机当且仅当你买机票了"

若 p 和 q 均为真或均为假，此语句为真。也就是说，如果你买机票了就能坐飞机，或是如果你没买机票就不能坐飞机。当 p 和 q 有对立的真值时，此语句为假，就是说，当你没买机票但却能坐飞机时(比如你获得一次免费旅行)或当你买了机票却不能坐飞机时(比如航空公司拒绝你登机)。 ■

在双蕴含中使用的"当且仅当"结构在普通语言中很少用。反之，通常用"如果，那么"或"仅当"结构来表示双蕴含。"当且仅当"的另一部分是隐含的。例如，考虑这个自然语言语句"你吃完饭才可以吃餐后甜点"。它意味着"你可以吃餐后甜点当且仅当你吃完饭"，这个语句逻辑等价于"如果你吃完饭，那么你可以吃餐后甜点"和"你可以吃餐后甜点，仅当你吃完饭"。后面这个语句在逻辑上等价于两个语句，一个是"如果你吃完饭，那么你可以吃甜点"，另一个是"仅当你吃完了饭，你才能吃甜点。"由于自然语言的这种不精确性，我们需要对自然语言中的

条件语句是否隐含它的逆作出假设。因为数学和逻辑注重精确，所以我们总是区分蕴含语句 $p \to q$和双蕴含语句 $p \leftrightarrow q$。

1.1.4 逻辑运算符的优先级

使用目前为止学到的否定运算符和逻辑运算符，可以构造复合命题。一般地，我们将使用括号来指明复合命题中的逻辑运算符的顺序。例如，$(p \vee q) \wedge (\neg r)$是 $p \vee q$ 和 $\neg r$ 的合取。然而，为了减少括号的数量，我们指定否定运算符在其他所有逻辑运算符前执行。这意味着$\neg p \wedge q$ 是 $\neg p$ 和 q 的合取，即$(\neg p) \wedge q$，而不是 p 和 q 的合取的否定，即 $\neg (p \wedge q)$。

另一个常用的优先级规则是合取运算符优先于析取运算符，因此 $p \wedge q \vee r$ 意味着$(p \wedge q) \vee r$，而不是 $p \wedge (q \vee r)$。因为这个规则不太好记，所以我们将继续使用括号来区别析取运算符和合取运算符的顺序。

最后，一个被接受的规则是条件运算符和双蕴含运算符的优先级低于合取和析取运算符。因此，$p \vee q \to r$ 等同于$(p \vee q) \to r$。当涉及条件运算符和双蕴含运算符时，我们也将使用括号，尽管条件运算的优先级高于双蕴含运算。表 1-7 示出各种逻辑运算符的优先级。

表 1-7 逻辑运算符的优先级

运算符	优先级
$\neg$	1
$\wedge$	2
$\vee$	3
$\to$	4
$\leftrightarrow$	5

1.1.5 翻译语言的句子

Extra Examples

有许多理由需把语言翻译成由命题变量和逻辑联接词组成的表达式，特别是因为语言（包括一切人类语言）常有二义性，把句子译成逻辑表达式可以消除歧义。注意，做这种翻译也许要在句子含义的基础上做些合理的假设。此外，一旦我们完成了从句子到逻辑表达式的翻译，我们就可以分析这些逻辑表达式以决定它们的真值。我们还可以对它们进行处理，并用推理规则（参见 1.5 节）对它们进行推理。

为说明把语句翻译成逻辑表达式的过程，考虑下面两个例子。

例 9 怎样把下面的句子翻译成逻辑表达式？

“只有你主修计算机科学或不是新生，才可以从校园网访问因特网。”

解 有许多方法翻译这一句子为逻辑表达式。尽管可以用一个命题变量，例如 p 来表示这一句子，但在分析其含义或用它做推理时，这种表示不会有什么作用。我们的办法是用命题变量表示其中的每一个句子成分，并找出其间合适的逻辑联接词。具体地说，令 a、c 和 f 分别表示“你可以从校园网访问因特网”、“你主修计算机科学”和“你是个新生”。注意到“只有……才”是表达蕴含的一种方式，上述句子可以译为

$$a \to (c \vee \neg f)$$

■

例 10 怎样把下面的句子翻译成逻辑表达式？

“除非你已满 16 周岁，否则只要你身高不足 4 英尺[⊖]就不能乘公园滑行铁道游乐车。”

解 有许多方式把这一句子翻译成逻辑表达式。最简单也最无用的方式是用一个命题变量，例如 p 表示这一句子。尽管这样做并不错，但当我们尝试分析这一句子，或用它做推理时，这种翻译对我们不会有什么帮助。较合适的方法是用命题变量表示其中的每一个句子成分，再找出它们之间合适的逻辑联接词。具体地说，令 p、r 和 s 分别表示“你能乘公园滑行铁道游乐车”、

⊖ 1 英尺＝30.48 厘米。——译者注

“你身高不足 4 英尺”和“你已满 16 周岁”，则上述句子可以翻译为

$$(r \wedge \neg s) \rightarrow \neg q$$

当然，还有其他方式可以把上述句子表示为逻辑表达式，但上面我们使用的这一表达式已满足我们的需要。■

1.1.6 系统规范说明

在说明硬件系统和软件系统时，将自然语言语句翻译成逻辑表达式是很重要的一部分。系统和软件工程师从自然语言中提取需求，生成精确、无二义性的规范说明，这些规范说明可作为系统开发的基础。例 11 说明如何在这一过程中使用命题表达式。

Extra Examples **例 11** 使用逻辑联接词表示规范说明“当文件系统满时，自动应答不能够发出”。

解 翻译这个规范说明的方法之一是令 p 表示“自动应答能够发出”，令 q 表示“文件系统满了”。则 $\neg p$ 表示“自动应答不能够发出”。因此，题中的规范说明可以用蕴含 $q \rightarrow \neg p$ 来表示。■

系统规范说明不应该包含有冲突的需求。若有，则无法开发出满足所有规范说明的系统。因此，表示这些规范说明的命题表达式应该是一致的。也就是说，对表达式中各个变量，必然有一个真值赋值使所有表达式为真。

例 12 确定下列系统规范说明是否一致：

“诊断消息存储在缓冲区中或是被重传。”

“诊断消息没有存储在缓冲区中。”

“如果诊断消息存储在缓冲区中，那么它被重传。”

Extra Examples **解** 要判断这些规范说明是否一致，我们首先用逻辑表达式表示它们。令 p 为“诊断消息存储在缓冲区中”，令 q 表示“诊断消息被重传”。则上面几个规范说明可以写为 $p \vee q$，$\neg p$ 和 $p \rightarrow q$。使所有三个规范说明为真的一个真值赋值必须有 p 为假使 $\neg p$ 为真。因为我们要使 $p \vee q$ 为真，但 p 必须为假，q 必须为真。由于当 p 为假且 q 为真时，$p \rightarrow q$ 为真，我们得出结论这些规范说明是一致的，因为当 p 为假且 q 为真时它们都是真的。使用一个真值表检验有关 p 和 q 的真值的四种可能赋值，我们可以得出同样的结论。■

例 13 如果在例 12 中加上一个系统规范说明“诊断消息不被重传”，它们还能保持一致吗？

解 由例 12 可知，只有当 p 为假且 q 为真时那三个规范说明才为真。然而，本例中的新规范说明是 $\neg q$，当 q 为真时它为假。因此，这四个规范说明不是一致的。■

1.1.7 布尔检索

Links 逻辑联接词广泛用于在大量信息检索中，例如，检索网页索引。由于这些检索使用来自命题逻辑的技术，所以称为布尔检索。

Extra Examples 在布尔检索中，联接词 *AND* 用于匹配包含两个检索项的记录，联接词 *OR* 用于匹配两个检索项之一或两项均匹配的记录，而联接词 *NOT*（有时写作 *AND NOT*）用于排除某个特定的检索项。当用布尔检索为有潜在价值的信息定位时，常需要细心安排逻辑联接词的使用。下面的例子说明布尔检索是怎样执行的。

例 14 **网页检索**。大部分网上检索引擎支持布尔检索技术，它有助于找到有关特定主题的网页。例如，要用布尔检索找出关于新墨西哥州（New Mexico）各大学的网页，我们可以寻找与 NEW *AND* MEXICO *AND* UNIVERSITIES 匹配的网页，检索的结果将包括含 NEW（新）、

MEXICO(墨西哥)和 UNIVERSITIES(大学)三个词的那些网页。这里面包含了所有我们感兴趣的网页，还包括其他网页，例如有关墨西哥国的新大学的网页。另一例子是，要找出与新墨西哥州或亚利桑那州(Arizona)的大学有关的网页，我们可以检索与(NEW *AND* MEXICO *OR* ARIZONA)*AND* UNIVERSITIES 匹配的网页。(注意，其中联接词 *AND* 优先于联接词 *OR*。)这一检索的结果将包括含 UNIVERSITIES 一词和 NEW 与 MEXICO 两词的所有网页及含 UNIVERSITIES 一词和 ARIZONA 一词的所有网页。同样除这两类我们感兴趣的网页外还会列出其他网页。最后，要想找出有关墨西哥国(不是新墨西哥州)的大学的网页，可以先找与 MEXICO *AND* UNIVERSITIES 匹配的网页，但由于这一检索的结果将会包括有关新墨西哥州的大学的网页以及墨西哥国的大学的网页，所以更好的办法是检索与(MEXICO *AND* UNIVERSITIES) *NOT* NEW 匹配的网页。这一检索的结果将包括含 MEXICO 和 UNIVERSITIES 两个词但不含词 NEW 的所有网页。 ■

1.1.8 逻辑难题

可以用逻辑推理解决的难题称为逻辑难题。求解逻辑难题是实践逻辑规则的一种好方法。同样，设计用于执行逻辑推理的计算机程序通常也使用著名的逻辑难题来演示它们的能力。许多人对求解逻辑难题感兴趣，有许多书和杂志也登载逻辑难题以供娱乐。

在此我们将讨论两个逻辑难题。首先介绍一个由逻辑学家和难题制作大师 Raymond Smullyan[⊖] 提出的难题，Smullyan 已经出版了十多本与逻辑难题(涉及逻辑推理)有关的书。

例 15 Smullyan[Sm78]中提出了许多与如下情形有关的难题：一个岛上居住着两类人——骑士和流氓。骑士说的都是实话，而流氓只会说谎。你碰到两个人 A 和 B，如果 A 说"B 是骑士"，B 说"我们两人不是一类人"。请判断 A、B 两人到底是流氓还是骑士。

解 令 p 和 q 分别表示命题 A 是骑士和 B 是骑士，则 $\neg p$ 和 $\neg q$ 分别表示 A 是流氓和 B 是流氓。

我们首先考虑 A 是骑士的情形，这就是说 p 是真的。如果 A 是骑士，那他说"B 是骑士"就是真话，因此 q 为真，A 和 B 就是一类人。然而，如果 B 是骑士，那么 B 说的话应该为真，然而却并非如此，因为 A 和 B 都是骑士。因此，我们可以得出 A 不是骑士，也就是，p 为假。

如果 A 是流氓，则据题目可知 A 所说的"B 是骑士"就不是真话，这意味着 q 为假且 B 也是流氓。而且，如果 B 是流氓，那么 B 说的话也是假的，这与 A 和 B 都是流氓是一致的。所以，我们得出结论 A 和 B 都是流氓。 ■

⊖ **Raymond Smullyan(1919 年生)**。Smullyan 中学就辍学了，他想学自己感兴趣的东西而不是中学课本上的知识。在换了一所大学后，他于 1955 年在芝加哥大学数学专业大学毕业。他靠在社团组织和俱乐部里表演魔术来赚取大学学费。1959 年，他于普林斯顿大学获得了逻辑学博士学位，师从 Alonzo Church。之后，Smullyan 陆续在达特茅斯学院、普林斯顿大学、犹太大学和纽约城市大学教授数学和逻辑学。1981 年，他加入印第安那大学哲学系，现在他是该校的名誉退休教授。

Smullyan 写过许多关于娱乐逻辑和娱乐数学的书，包括 *Satan, Cantor, and Infinity*；*What Is the Name of This Book?*；*The Lady or the Tiger?*；*Alice in Puzzleland*；*To Mock a Mockingbird*；*Forever Undecided*；以及 *The Riddle of Scheherazade: Amazing Logic Puzzles, Ancient and Modern*。因为他提出逻辑难题很有挑战性和娱乐性并能激发思维，他被认为是现代的 Lewis Carroll。Smullyan 也写过几本关于国际象棋推理逻辑应用方面的书，以及几本高级的关于数学逻辑和集合论的书。Smullyan 对自引用(self-reference)特别感兴趣，并且乐衷于向大众解释数学逻辑思想。

Smullyan 是一位天才的音乐家，经常与妻子一同弹奏钢琴，妻子是一位专业的钢琴演奏家。制造望远镜是他的一个嗜好。他还对光学和立体照相感兴趣。他曾说"我从不像某些人那样不能同时兼顾教学和研究，因为我在教学的同时就在做研究。"

在本节末的练习 51～55 中，我们会进一步讨论 Smullyan 的骑士和流氓的难题。接下来，我们介绍一个与两个孩子有关的“泥巴孩子难题”(muddy children puzzle)。

例 16 父亲让两个孩子(一个男孩，一个女孩)在后院玩耍，并让他们不要把身上搞脏。然而，在玩的过程中，两个孩子都在额头上沾了泥。当孩子们回来后，父亲说“你们当中至少有一个人额头上有泥”，然后他问每一个孩子“你知道你额头上有没有泥吗?”同样的问题对每个孩子都问了两遍。假设每个孩子都可以看到对方的额头上是否有泥，但不能看见自己的额头，孩子们将会怎样回答呢? 假设两个孩子都很诚实并且都同时回答每一次提问。

解 令 s 和 d 分别表示命题儿子的额头上有泥和女儿的额头上有泥。当父亲说“你们当中至少有一个人额头上有泥”时，表示的是 $s \vee d$ 为真。当父亲第一次问那个问题时两个孩子都将回答“不知道”，因为他们都看到对方的额头上有泥。也就是说，儿子知道 d 为真，但不知道 s 是否为真。而女儿知道 s 为真，但不知道 d 是否为真。

在儿子对第一次询问回答“不知道”后，女儿可以判断出 d 必为真。这是因为问第一次问题时，儿子知道 $s \vee d$ 为真，但不能判断 s 是否为真。利用这个信息，女儿能够得出结论 d 必定为真，因为如果 d 为假，儿子有理由推出，由于 $s \vee d$ 为真，那么 s 必定为真，因此他对第一个问题的回答应为“知道”。儿子也可以类似推断出 s 必为真。因此，第二次两个孩子都将回答“知道”。■

1.1.9 逻辑运算和位运算



真值	位
T	1
F	0

计算机用字位表示信息，每个字位有两个可能的值，即 0 或 1。字位的这一含义来自二进制数字，因为 0 和 1 是用在数的二进制表示中的数字[⊖]。1946 年著名的统计学家图凯(John TuKey)[⊜]引入了这一术语。字位可以用于表示真值，因为只有两个真值，即真与假。习惯上用 1 表示真，用 0 表示假。就是 1 表示 T(真)，0 表示 F(假)。若变量的值或为真或为假，则此变量称为*布尔变量*。于是一个布尔变量可以用一个字位表示。

计算机的*字位运算*对应于逻辑联接词，只要在运算符 ∧、∨ 和 ⊕ 的真值表中用 1 代替 T，用 0 代替 F，就能得到表 1-8 所示的对应的字位运算表。我们还将用符号 OR、AND 和 XOR 表示运算符 ∨、∧ 和 ⊕，许多程序设计语言正是这样表示的。

信息一般用位串表示，也就是用 0 和 1 的序列表示。这时，对位串的运算即可用来处理信息。

⊖ 历史注记：有人建议过用别的词称呼二进制数字，例如，binit 和 bigit，但从来没有被广泛地接受。采用 bit 是因为它作为英语常用词的含义。要了解图凯选用 bit 一词的报道，参见 1984 年 4 月期的《*Annals of the History of Computing*》。



⊜ **图凯(John Wilder Tukey, 1915—2000)**生于马萨诸塞州新贝德福德，是个独生子。他的双亲都是教员，认为家庭教育最适合开发他的潜力。他的正规教育从布朗大学开始，学的是数学和化学。他在布朗大学获得化学硕士学位，接着在普林斯顿大学继续学习，不过从化学改学数学。1939 年由于他在拓扑学方面的工作，获得普林斯顿大学博士学位，同时还被任命为普林斯顿大学数学教师。随着第二次世界大战的爆发，他参加了火力控制研究办公室(Fire Control Research Office)的工作，开始研究统计学。图凯发现统计研究很合他的口味，并以他的技能给几位最有影响的统计学家留下了深刻印象。1945 年二战结束时，图凯回到普林斯顿大学数学系担任统计学教授，并在 AT&T 贝尔实验室兼职。图凯于 1966 年创立了普林斯顿大学统计学系，成为该系首任主任。图凯在统计学的许多领域作出了重要贡献，包括方差分析、时间序列谱评估、关于采自一次试验的一组参数值的推断以及统计学原理。不过他最著名的工作是他与库雷(J. W. Cooley)共同发明的快速傅里叶变换。

服务于总统科学顾问委员会期间，图凯以他的洞察力和经验作出了重要贡献。他曾担任过处理环境、教育以及化学与健康等事务的好几个重要委员会的主席。他还参与了几个核裁军委员会的工作。图凯得过许多奖，包括国家科学奖章。

表 1-8　字位运算符 OR、AND 和 XOR 的真值表

x	y	$x \vee y$	$x \wedge y$	$x \oplus y$
0	0	0	0	0
0	1	1	0	1
1	0	1	0	1
1	1	1	1	0

定义 7　位串是 0 个或多个字位的序列。位串的长度就是它所含的字位的个数。

例 17　101010011 是长度为 9 的一个位串。■

我们可以把字位运算扩展到位串上。我们在长度相同的两个位串上定义它们的按位 OR、按位 AND 和按位 XOR 分别为这样的位串，其中每个字位均由被运算的两个位串对应字位经 OR、AND 和 XOR 运算而得。我们分别用符号 ∨、∧ 和⊕表示按位 OR、按位 AND 和按位 XOR 运算。我们用下面的例子解释对位串的这些按位运算。

例 18　求位串 01 1011 0110 和 11 0001 1101 的按位 OR、按位 AND 和按位 XOR(这里以及本书其他地方均把位串按 4 位分块以便于阅读)。

解　这两个位串的按位 OR、按位 AND 和按位 XOR 分别由对应字位的 OR、AND 和 XOR 得到，其结果是

01 1011 0110
11 0001 1101
11 1011 1111　按位 OR
01 0001 0100　按位 AND
10 1010 1011　按位 XOR

■

练习

1. 下列哪些语句是命题？这些是命题的语句的真值是什么？
   a)波士顿是马萨诸塞州首府。
   b)迈阿密是佛罗里达州首府。
   c)$2+3=5$。
   d)$5+7=10$。
   e)$x+2=11$。
   f)回答这一问题。
   g)对每一对实数 x、y，都有 $x+y=y+x$。
2. 下列哪些是命题？这些命题的真值是什么？
   a)别过去。
   b)几点了？
   c)在缅因州没有黑苍蝇。
   d)$4+x=5$。
   e)若 $x=1$，则 $x+1=5$。
   f)若 $x=z$，则 $x+y=y+z$。
3. 下列各命题的否定是什么？
   a)今天是星期四。
   b)新泽西没有污染。
   c)$2+1=3$。
   d)缅因州的夏天又热又晒。
4. 令 p、q 为如下命题：
   p：本周我买了一张彩票。
   q：星期五我中了百万美元头奖。
   把下列各命题表达为汉语句子：
   a)$\neg p$　b)$p \vee q$　c)$p \rightarrow q$　d)$p \wedge q$
   e)$p \leftrightarrow q$　f)$\neg p \rightarrow \neg q$　g)$\neg p \wedge \neg q$　h)$\neg p \vee (p \wedge q)$

5. 令 p 和 q 分别表示命题"在新泽西海岸游泳是允许的"和"在海岸附近发现过鲨鱼"。把下列各命题表达为汉语句子：

a) $\neg q$　b) $p \land q$　c) $\neg p \lor q$　d) $p \to \neg q$

e) $\neg q \to p$　f) $\neg p \to \neg q$　g) $p \leftrightarrow \neg q$　h) $\neg p \land (p \lor \neg q)$

6. 令 p 和 q 分别表示命题"选举已经有了结果"和"选票已经计数完毕"。把下列各命题表达为汉语句子：

a) $\neg p$　b) $p \lor q$　c) $\neg p \land q$　d) $q \to p$

e) $\neg q \to \neg p$　f) $\neg p \to \neg q$　g) $p \leftrightarrow q$　h) $\neg q \lor (\neg p \land q)$

7. 令 p、q 为如下命题：

p：气温在零度以下。

q：正在下雪。

用 p、q 和逻辑联接词写出下列各命题：

a)气温在零度以下且正在下雪。

b)气温在零度以下，但没下雪。

c)气温不在零度以下，也不下雪。

d)也许在下雪，也许在零度以下(也许两者在内)。

e)若气温在零度以下，那也就在下雪。

f)也许气温在零度以下，也许在下雪，但如果在零度以下，就不在下雪。

g)气温在零度以下是下雪的充分必要条件。

8. 令 p、q 和 r 为如下命题：

p：你得流感了。

q：你错过了期终的考试。

r：这门课你及格了。

将下列各命题用汉语表示：

a) $p \to q$　b) $\neg q \leftrightarrow r$　c) $q \to \neg r$

d) $p \lor q \lor r$　e) $(p \to \neg r) \lor (q \to \neg r)$　f) $(p \land q) \lor (\neg q \land r)$

9. 令 p、q 为如下命题：

p：你的车速超过每小时 65 英里(1 英里＝1.6 公里)。

q：你接到一张超速罚单。

用 p、q 和逻辑联接词写出下列命题：

a)你的车速没有超过每小时 65 英里。

b)你的车速超过每小时 65 英里，但没接到超速罚单。

c)你的车速若超过每小时 65 英里，将接到一张超速罚单。

d)你的车速不超过每小时 65 英里，就不会接到超速罚单。

e)车速超过每小时 65 英里足以接到超速罚单。

f)你接到一张超速罚单，但你的车速没超过每小时 65 英里。

g)只要你接到一张超速罚单，你的车速就超过每小时 65 英里。

10. 令 p、q、r 为如下命题：

p：你的期末考试得了个 A。

q：你做了本书每一道练习。

r：这门课你得了个 A。

用 p、q、r 和逻辑联接词写出下列命题：

a)这门课你得了个 A，但你并没做本书的每道练习。

b)你的期末考试得了个 A，你做了本书的每一道练习，并且这门课你得了个 A。

c)想在这门课得 A，你必须在期末考试得 A。

d)你的期末考试得了个 A，你没有做本书的每道练习，可不管怎样这门课你得了个 A。

e)期末考试得 A 并且做本书的每道练习，足以使你这门课得 A。

f)这门课得 A 当且仅当你做本书的每道练习或期末考试得 A。

11. 令 p、q、r 为如下命题：

p：在这个地区发现过灰熊。

q：在小路上徒步旅行是安全的。

r：小路两旁的草莓成熟了。

用 p、q、r 和逻辑联接词写出下列命题：

a)小路两旁的草莓成熟了，但在这个地区没有发现过灰熊。

b)在这个地区没有发现过灰熊，且在小路上徒步旅行是安全的，但小路两旁的草莓成熟了。

c)如果小路两旁的草莓成熟了，徒步旅行是安全的当且仅当在这个地区没有发现过灰熊。

d)在小路上徒步旅行不安全，但在这个地区没有发现过灰熊且小路两旁的草莓成熟了。

e)为了使在小路上旅行很安全，其必要条件(但并不充分)是小路两旁的草莓没有成熟且在这个地区没有发现过灰熊。

f)无论何时在这个地区发现过灰熊且小路两旁的草莓成熟了，在小路上徒步旅行就不安全。

12. 判断下列这些双蕴含是真是假：

a)$2+2=4$ 当且仅当 $1+1=2$。

b)$1+1=2$ 当且仅当 $2+3=4$。

c)现在是冬季当且仅当现在不是春季、夏季或秋季。

d)$1+1=3$ 当且仅当猪会飞。

e)$0>1$ 当且仅当 $2>1$。

13. 判断下列各蕴含是真是假：

a)若 $1+1=2$，则 $2+2=5$。

b)若 $1+1=3$，则 $2+2=4$。

c)若 $1+1=3$，则 $2+2=5$。

d)若猪会飞，那么 $1+1=3$。

e)若 $1+1=3$，就存在上帝。

f)若 $1+1=3$，猪就会飞。

g)若 $1+1=2$，猪就会飞。

h)若 $2+2=4$，则 $1+2=3$。

14. 就下列各语句，判断其中想表达的是同或还是异或，说明理由：

a)要求有使用 C ++ 或 Java 的经验。

b)午餐包括汤或沙拉。

c)你必须持护照或选民登记卡才能入境。

d)出版或销毁。

15. 对下列各语句，说一说其中的或是同或(即析取)与异或时它们的含义。你认为语句想表示的是哪个或？

a)要选修离散数学课，你必须已经选修微积分或计算机科学的一门课。

b)从 Acme 汽车公司购买一部新车，你就能得到 2 000 美元现金回扣，或 2%的汽车贷款。

c)两人套餐包括 A 列中的两项或 B 列中的三项。

d)若下雪超过两英尺或寒风低于－100，学校就停课。

16. 把下列语句写成“如果 p，那么 q”的形式。[提示：参考本节列出的通常表达蕴含的方式。]

a)要想晋升，帮老板洗车是很有必要的。

b)吹南风预示着春天要来了。

c)保单有效的充分条件是你的计算机购买时间不超过一年。

d)Willy 只要行骗就会被抓住。

e)只有支付了订阅费，你才能访问网站。

f)获得选举源于认识右翼人士。

g)每当坐船 Carol 就会晕船。

17. 把下列语句写成“如果 p，那么 q”的形式。[提示：参考本节列出的通常表达蕴含的方式。]

a)吹东北风的时候就下雪。

b)暖天持续一周苹果树就开花。

c)活塞队赢得冠军就意味着他们打败了湖人队。

d)必须走 8 英里才能到朗斯峰的顶峰。

e)要得到教授职位，世界闻名就够了。

f)如果你驾车超过 400 英里，就需要买汽油了。

g)只有你购买的 CD 机不超过 90 天，你的保修单才有效。

18. 把下列语句写成“如果 p，那么 q”的形式。[提示：参考本节列出的通常表达蕴含的方式。]

a)只要你发给我一个电子邮件，我就会记住把地址寄给你。

b)要成为美国公民，只要你生在美国就行了。

c)如果你保存课本，它会是你未来其他课程的有用参考书。

d)如果守门员表现好，红翼队将赢得斯坦利杯。

e)你获得这一职位表明你有最好的信誉。

f)有风暴时沙滩受侵蚀。

g)在服务器登录必须有一个有效的口令。

19. 把下列命题写成“p 当且仅当 q”的形式：

a)如果外边热你就买冰激淋；如果你买冰激淋，外边就热。

b)你赢得竞赛的充分必要条件是你有那唯一的获胜券。

c)只有你有关系才能得到提拔，只有得到提拔你才有关系。

d)如果你看电视，记忆会衰退；反之亦然。

e)火车恰在我乘坐的那些日子晚点。

20. 把下列命题写成“p 当且仅当 q”的形式：

a)你能在这门课得 A 的充分必要条件是学会解离散数学问题。

b)如果你每天看报，你就了解情况；反之亦然。

c)周末天下雨，下雨天是周末。

d)仅当巫师不在家时你能看到他，仅当你能看到巫师时他不在家。

21. 给出下列各蕴含关系的逆、逆否和否命题：

a)如果今天下雪，我明天就去滑雪。

b)只要有测验，我就来上课。

c)只有当正整数没有 1 和它自己以外的因数时，它才是素数。

22. 给出下列蕴含关系的逆、逆否和否命题：

a)如果今晚下雪，我将呆在家里。

b)只要是阳光充足的夏天，我就去海滩。

c)如果我起床晚了，一定是我白天睡到了中午。

23. 为下列各复合命题构造真值表：

a) $p \wedge \neg p$　　b) $p \vee \neg p$　　c) $(p \vee \neg q) \rightarrow q$

d) $(p \vee q) \rightarrow (p \wedge q)$　　e) $(p \rightarrow q) \leftrightarrow (\neg q \rightarrow \neg p)$　　f) $(p \rightarrow q) \rightarrow (q \rightarrow p)$

24. 为下列各复合命题构造真值表：

a) $p \rightarrow \neg p$　　b) $p \leftrightarrow \neg p$　　c) $p \oplus (p \vee q)$

d) $(p \wedge q) \rightarrow (p \vee q)$　　e) $(q \rightarrow \neg p) \leftrightarrow (p \leftrightarrow q)$　　f) $(p \leftrightarrow q) \oplus (p \leftrightarrow \neg q)$

25. 为下列各复合命题构造真值表：

a) $(p \vee q) \rightarrow (p \oplus q)$　　b) $(p \oplus q) \rightarrow (p \wedge q)$　　c) $(p \vee q) \oplus (p \wedge q)$

d) $(p \leftrightarrow q) \oplus (\neg p \leftrightarrow q)$　　e) $(p \leftrightarrow q) \oplus (\neg p \leftrightarrow \neg r)$　　f) $(p \oplus q) \rightarrow (p \oplus \neg q)$

26. 为下列各复合命题构造真值表：

a) $p \oplus p$　　b) $p \oplus \neg p$　　c) $p \oplus \neg q$

d) $\neg p \oplus \neg q$　　e) $(p \oplus q) \vee (p \oplus \neg q)$　　f) $(p \oplus q) \wedge (p \oplus \neg q)$

27. 为下列各复合命题构造真值表：

a) $p \to \neg q$　　b) $\neg p \leftrightarrow q$　　c) $(p \to q) \vee (\neg p \to q)$

d) $(p \to q) \wedge (\neg p \to q)$　　e) $(p \leftrightarrow q) \vee (\neg p \leftrightarrow q)$　　f) $(\neg p \leftrightarrow \neg q) \leftrightarrow (p \leftrightarrow q)$

28. 为下列各复合命题构造真值表：

a) $(p \vee q) \vee r$　　b) $(p \vee q) \wedge r$　　c) $(p \wedge q) \vee r$

d) $(p \vee q) \wedge r$　　e) $(p \vee q) \wedge \neg r$　　f) $(p \wedge q) \vee \neg r$

29. 为下列各复合命题构造真值表：

a) $p \to (\neg q \vee r)$　　b) $\neg p \to (q \to r)$　　c) $(p \to q) \vee (\neg p \to r)$

d) $(p \to q) \wedge (\neg p \to r)$　　e) $(p \leftrightarrow q) \vee (\neg q \leftrightarrow r)$　　f) $(\neg p \leftrightarrow \neg q) \leftrightarrow (q \leftrightarrow r)$

30. 构造 $((p \to q) \to r) \to s$ 的真值表。

31. 构造 $(p \leftrightarrow q) \leftrightarrow (r \leftrightarrow s)$ 的真值表。

32. 假定在计算机程序中，执行下列语句之前，$x=1$，在执行之后 x 的值是什么？

a) **if** $1+2=3$ **then** $x := x+1$

b) **if** $(1+1=3)$ *OR* $(2+2=3)$ **then** $x := x+1$

c) **if** $(2+3=5)$ *AND* $(3+4=7)$ **then** $x := x+1$

d) **if** $(1+1=2)$ *XOR* $(1+2=3)$ **then** $x := x+1$

e) **if** $x<2$ **then** $x := x+1$

33. 求下列各对位串的按位 *OR*、按位 *AND* 及按位 *XOR*：

a) 101 1110，010 0001　　b) 1111 0000，1010 1010

c) 00 0111 0001，10 0100 1000　　d) 11 1111 1111，00 0000 0000

34. 计算下列表达式：

a) $1\ 1000 \wedge (0\ 1011 \vee 1\ 1011)$　　b) $(0\ 1111 \wedge 1\ 0101) \vee 0\ 1000$

c) $(0\ 1010 \oplus 1\ 1011) \oplus 0\ 1000$　　d) $(1\ 1011 \vee 0\ 1010) \wedge (1\ 0001 \vee 1\ 1011)$



模糊逻辑可用于人工智能。在模糊逻辑中命题的真值是界于 0 和 1(包括 0 和 1)之间的数。以 0 为真值的命题为假，以 1 为真值的命题为真。0 和 1 之间的真值表示不同程度的真值。例如，语句"傅雷德是幸福的"的真值可以是 0.8，因为傅雷德大部分时间是幸福的；"约翰是幸福的"的真值可能是 0.4，因为他幸福的时间比一半稍短。

35. 模糊逻辑中命题否定的真值是 1 减去该命题的真值。语句"傅雷德不幸福"和"约翰不幸福"的真值是什么？

36. 模糊逻辑中两个命题的合取的真值是两个命题真值的最小值。语句"傅雷德和约翰都幸福"和"傅雷德和约翰都不幸福"的真值是什么？

37. 模糊逻辑中两个命题的析取的真值是两个命题真值的最大值。语句"傅雷德幸福或约翰幸福"与"傅雷德不幸福或约翰不幸福"的真值是什么？

*38. 断言"本语句为假"是命题吗？

*39. 在一个 100 条语句的列表中，第 n 条语句是"在这个列表中，恰有 n 个语句为假。"

a) 从这些语句中你可以得出什么结论？

b) 若第 n 条语句是"在这个列表中，至少有 n 个语句为假"，回答问题 a)。

c) 假设这个列表包含 99 条语句，回答问题 b)。

40. 在古西西里的传说中有一个住在边远小镇上的剃头匠，只有穿过一条危险的山路才能找到他。这个剃头匠只给那些不自己剃须的人刮胡子。这样的剃头匠能存在吗？

41. 边远村庄的每个人要么总说真话，要么总说谎。对旅游者的问题，村民要么回答"是"，要么回答"不"。假定你在这一地区旅游，走到了一个岔路口，一条岔路通向你想去的遗址，另一岔路通向丛林深处。此时恰有一村民站在岔路口，问村民什么样的一个问题就能决定走哪条路？

42. 一个探险者被几个吃人者抓住了。有两种吃人者：总是说谎的和永不说谎的。除非探险者能判断出一位指定的吃人者是说谎者还是说真话者，否则就要被吃人者烤了吃。探险者只被允许问这位吃人者一个问题。

a)解释为什么问："你说谎吗?"是不行的。

b)找一个问题，使探险者可以用来判断该吃人者是说谎者还是说真话者。

43. 使用命题 p"扫描消息中的病毒"和 q"消息来自一个未知的系统"以及逻辑联接词来表达下列系统规范说明。

a)"每当消息来自一个未知的系统时，就扫描消息中的病毒。"

b)"消息来自一个未知的系统，但不扫描消息中的病毒。"

c)"每当消息来自一个未知的系统时，就有必要扫描消息中的病毒。"

d)"当消息不是来自一个未知的系统时，就不扫描消息中的病毒。"

44. 使用命题 p"用户输入有效的口令"、q"访问被授权"、r"用户已经支付了订阅费"以及逻辑联接词来表达下列系统规范说明。

a)"用户已经支付了订阅费，但没有输入有效的口令。"

b)"每当用户已经支付了订阅费并输入有效的口令，访问被授权。"

c)"如果用户没有支付订阅费，访问被拒绝。"

d)"如果用户没有输入有效的口令但已经支付了订阅费，访问被授权。"

45. 下列系统规范说明一致吗?

"当且仅当系统正常操作时，系统处于多用户状态。如果系统正常操作，则它的核心程序正在运行。核心程序不能正常运行，或者系统处于中断模式。如果系统不处于多用户状态，它就处于中断模式。系统不处在中断模式。"

46. 下列系统规范说明一致吗?

"每当对系统软件进行升级时，用户不能访问文件系统。如果用户能访问文件系统，那么他们能保存新文件。如果用户不能保存新文件，那么系统软件尚未升级。"

47. 下列系统规范说明一致吗?

"路由器能向边缘系统发送分组仅当它支持新的地址空间时。要让路由器支持新的地址空间，就必须安装最新的软件发布。如果最新的软件发布被安装了，路由器就能向边缘系统发送分组。路由器不支持新的地址空间。"

48. 下列系统规范说明一致吗?

"如果文件系统未加锁，那么新消息将被排成队。如果文件系统未加锁，则系统正常运行；反之亦然。如果新消息尚未排队，就会送入消息缓冲区。如果文件系统未加锁，那么新消息将被送入消息缓冲区。新消息不会被送入消息缓冲区。"

49. 你会用什么样的布尔检索来寻找关于新泽西州海滩的网页?如果你想找关于泽西岛(在英吉利海峡)海滩的网页呢?

50. 你会用什么样的布尔检索来寻找关于徒步旅行西弗吉尼亚的网页?如果你想找关于徒步旅行弗吉尼亚的网页，而不是西弗吉尼亚呢?

练习 51～55 是关于 Smullyan 的趣味逻辑的。一个岛上居住着两类人：骑士和流氓。骑士说的都是实话，而流氓只会说谎。根据下列不同情况判断 A、B 两人到底是流氓还是骑士，如果认为得不出最终结论的，也请说出任何可能的结论。

51. A 说："我们之间至少有一个是流氓"，B 什么都没说。

52. A 说："我们两个都是骑士"，B 说："A 是流氓"。

53. A 说："我是流氓或者 B 是骑士"，B 什么都没说。

54. A 和 B 都说："我是骑士"。

55. A 说："我们都是流氓"，B 什么都没说。

练习 56～61 是智力游戏题，解题时可以先把语句翻译成逻辑表达式，再用真值表从这些表达式作推理。

56. 对于杀害 Cooper 先生的凶手，警察有三个怀疑对象：Smith 先生、Jones 先生和 Williams 先生。他们三人都声称没有杀害 Cooper。Smith 还称 Cooper 是 Jones 的朋友并且 Williams 不喜欢他。Jones 也声称他不认识

Cooper 并且 Cooper 被害的当天他不在镇上。Williams 也声称他在案发当天看见 Smith 和 Jones 与 Cooper 在一起，因此不是 Smith 就是 Jones 是凶手。在下列情况下，请判断谁是凶手。

a)三人中有一人是凶手，清白的那两个人说的是真话，但凶手说的话不一定为真。

b)清白者没有撒谎。

57. 斯蒂夫想用两个事实来判断三位工作伙伴的相对薪水。首先他知道如果傅雷德的薪水不是三人中最高的，那么杰尼斯的最高。其次他知道如果杰尼斯的薪水不是最低的，那么麦吉的最高。从以上斯蒂夫知道的事实，有可能决定傅雷德、麦吉和杰尼斯的相对薪水吗？如果能，谁的最高，谁的最低？解释你的推理。

58. 五个朋友都能进入谈话室。如果知道下面这些信息，能决定谁在谈话吗？凯文或希思或他们两个都在谈话。兰迪或维杰在谈话，但没有同时谈话。如果阿比在谈话，那么兰迪也在谈话。维杰和凯文或者两人都在谈话，或者都不谈话。如果希思在谈话，那么阿比和凯文也在谈。解释你的推理。

59. 侦探调查了罪案的四位证人。从证人的话侦探得出的结论是：如果男管家说的是真话，那么厨师说的也是真话；厨师和园丁说的不可能都是真话；园丁和杂役不可能都在说谎；如果杂役说真话，那么厨师在说谎。侦探能判定这四位证人分别是在说谎还是在说真话吗？解释你的推理。

60. 四个朋友被认定为非法进入某计算机系统的嫌疑人。他们已对调查员作了陈述。艾丽斯说“卡罗斯干的。”约翰说“我没干。”卡罗斯说“戴安娜干的。”戴安娜说“卡罗斯说是我干的，他说谎。”

a)如果调查员知道四个嫌疑人中恰有一人说真话，那么谁干的？解释你的推理。

b)如果调查员知道恰有一人说谎，谁干的？解释你的推理。

*61. 试求解下面这个由爱因斯坦提出的著名的逻辑难题(也称为斑马难题)。5 个不同国家且工作各不相同的人分别住在一条街上的 5 所房子里，每所房子的颜色不同，每个人都有自己养的不同的宠物，喜欢喝不同的饮料。根据以下提示，你能告诉我哪个房子里的人养斑马(zebra)，哪个房子里的人喜欢喝矿泉水吗？

英国人住在红色的房子里。西班牙人养了一条狗。日本人是一个油漆工。意大利人喜欢喝茶。挪威人住在左边的第一个房子里。绿房子在白房子的右边。摄影师养了一只蜗牛。外交官住在黄房子里。中间那个房子里的人喜欢喝牛奶。喜欢喝咖啡的人住在绿房子里。挪威人住在蓝色的房子旁边。小提琴家喜欢喝桔子汁。养狐狸的人所住的房子与医师的房子相邻。养马的人所住的房子与外交官的房子相邻。(提示：绘一张表，其中行表示每个人，列表示他们所住的房子的颜色、他们的工作、他们养的宠物以及他们喜欢喝的饮料，用逻辑推理来判断表中正确的项。)

1.2 命题等价

1.2.1 引言

数学证明中使用的一种重要方法是用真值相同的一个语句取代另一个语句。因此，在构造数学证明时广泛使用从已知复合命题产生具有同样真值的其他命题的方法。

我们就从根据可能的真值对复合命题进行分类开始讨论。

定义 1 复合命题称为*永真式*(或*重言式*)，如果无论其中出现的命题的真值是什么，它的真值总是真。真值永远为假的复合命题称为*矛盾*。最后，既不是永真式又不是矛盾的命题称为*可能式*。

在数学推理中永真式和矛盾往往是重要的，下面的例子解释了这两类命题。

例 1 我们可以只用一个命题构造永真式和矛盾。$p \vee \neg p$ 和 $p \wedge \neg p$ 的真值表如表 1-9 所示。因为 $p \vee \neg p$ 总是真，它是永真式。因为 $p \wedge \neg p$ 总是假，它是矛盾。 ■

表 1-9 永真式和矛盾的例子

p	$\neg p$	$p \vee \neg p$	$p \wedge \neg p$
T	F	T	F
F	T	T	F

1.2.2 逻辑等价

在所有可能的情况下都有相同真值的两个复合命题称为逻辑等价。我们也可以如下定义这一概念。

定义 2 如果 $p\leftrightarrow q$ 是永真式，命题 p 和 q 称为是逻辑等价的。记号 $p\equiv q$ 表示 p 和 q 逻辑等价。

注意 符号$\equiv$不是逻辑连接词，因为 $p\equiv q$ 不是复合命题，而语句 $p\leftrightarrow q$ 是永真式。符号$\Leftrightarrow$有时用来代替$\equiv$表示等价。

判定两个命题是否等价的方法之一是使用真值表。特别是，命题 p 和 q 等价当且仅当给出它们真值的两列完全一致。下面例子解释了这一方法。

例 2 证明$\neg(p\vee q)$和$\neg p\wedge\neg q$等价。这一等价关系是德摩根定律之一。这是用 19 世纪中叶英国数学家德摩根(Augustus De Morgan)[⊖]的名字命名的。

解 表 1-10 给出了这些命题的真值。由于对 p 和 q 所有可能的真值组合，命题$\neg(p\vee q)$和$\neg p\wedge\neg q$的真值都一样，这两个命题逻辑等价。

表 1-10 $\neg(p\vee q)$和$\neg p\wedge\neg q$的真值表

p	q	$p\vee q$	$\neg(p\vee q)$	$\neg p$	$\neg q$	$\neg p\wedge\neg q$
T	T	T	F	F	F	F
T	F	T	F	F	T	F
F	T	T	F	T	F	F
F	F	F	T	T	T	T

■

⊖ **奥古斯塔·德摩根(Augustus De Morgan, 1806—1871)**生于印度，他父亲是印度陆军上校。德摩根 7 个月大时全家移居英国。他上的是私立学校并在那里培养了对数学的兴趣。德摩根 1827 年毕业于剑桥三一学院。尽管他想过要学医或学法律，最后还是决定以数学为毕生事业。1828 年他获得了伦敦大学学院的一个职位，但当他的一位教授同事被无故解雇时他辞职了。不过，在 1836 年他的继任人去世后他又回到了自己的位置，直到 1866 年。

德摩根以强调原理胜于技术而著名。他的学生中有许多是著名的数学家，包括拉弗雷斯伯爵夫人奥古斯塔(Ada Augusta)，她是巴贝奇(Charles Babbage)计算机器研究的合作者(参见关于 Ada Augusta 的生平注释)。

德摩根是位特别多产的作家，他为不止 15 家期刊写了 1000 多篇文章。德摩根还为许多学科撰写课本，包括逻辑、概率、微积分和代数。1838 年，他首次给出了数学归纳法这一重要证明技术的清晰解释，数学归纳法这一术语即由他创造。18 世纪 40 年代他对符号逻辑的发展做出了奠基性的贡献。他发明了帮助他证明命题等价的符号，其中包括以他的名字命名的定律。1842 年，德摩根给出了关于极限的也许是第一个准确定义，并提出了无穷数列收敛的若干检验标准。德摩根还对数学史有兴趣，写了牛顿和哈雷的生平传记。

1837 年，德摩根与弗伦德(Sophia Frend)结婚，后者在 1882 年撰写了德摩根传记。德摩根的研究、写作和教学使他无暇顾及家庭和社交，不过他的善良、幽默及广博的知识仍是闻名于世的。

拉弗雷斯伯爵夫人奥古斯塔(Ada Augusta, 1815—1852) 艾达·奥古斯塔是著名诗人拜伦(Byron)勋爵和安娜贝纳·米尔班克(Annabella Millbanke)的唯一孩子，然而他们在艾达 1 个月大时就分居了。艾达由母亲养大，也是母亲发掘了艾达在智力上的天赋。艾达的老师是数学家威廉·弗伦德(William Frend)和奥古斯塔·德摩根(Augustus De Morgan)。1838 年，艾达同金(King)爵士结婚，此人后来被晋升为拉弗雷斯伯爵。他们一共有 3 个孩子。

艾达结婚后仍坚持她在数学上的研究，帮助查理·巴贝奇(Charles Babbage)从事早期的计算机器的研究工作，这种机器称为解析机。关于此种机器最完整的解释可在艾达的著作中找到。1845 年以后，艾达同巴贝奇一同致力于能预测赛马结果的系统的研究。然而，他们的系统不能运行，使艾达去世时还因此欠下大笔债务。程序设计语言 Ada 就是为纪念这位伯爵夫人而命名的。

例 3 证明命题 $p \to q$ 和 $\neg p \vee q$ 逻辑等价。

解 在表 1-11 中构造了这两个命题的真值表。由于 $\neg p \vee q$ 和 $p \to q$ 的真值一致，它们是逻辑等价的。

表 1-11 $\neg p \vee q$ 和 $p \to q$ 的真值表

p	q	$\neg p$	$\neg p \vee q$	$p \to q$
T	T	F	T	T
T	F	F	F	F
F	T	T	T	T
F	F	T	T	T

■

例 4 证明命题 $p \vee (q \wedge r)$ 和 $(p \vee q) \wedge (p \vee r)$ 逻辑等价。这是析取对合取的分配律。

解 表 1-12 中构造了这两个命题的真值表。因为 $p \vee (q \wedge r)$ 和 $(p \vee q) \wedge (p \vee r)$ 的真值一样，它们是逻辑等价的。

注意 涉及 3 个不同命题的复合命题的真值表需要 8 行，每一行代表这 3 个命题的真值的一种组合。一般涉及 n 个命题的复合命题需要 2^n 行。

表 1-12 $p \vee (q \wedge r)$ 和 $(p \vee q) \wedge (p \vee r)$ 逻辑等价的演示

p	q	r	$q \wedge r$	$p \vee (q \wedge r)$	$p \vee q$	$p \vee r$	$(p \vee q) \wedge (p \vee r)$
T	T	T	T	T	T	T	T
T	T	F	F	T	T	T	T
T	F	T	F	T	T	T	T
T	F	F	F	T	T	T	T
F	T	T	T	T	T	T	T
F	T	F	F	F	T	F	F
F	F	T	F	F	F	T	F
F	F	F	F	F	F	F	F

■

表 1-13 给出了若干重要的等价关系[⊖]。在这些等价关系中，**T** 表示永远为真的任何命题，**F** 表示永远为假的任何命题。对于涉及蕴含和双蕴含的复合命题，我们也在表 1-13～表 1-15 中给出了一些有用的等价关系。本节末的练习中要求读者证明这些等价关系。

析取的结合律表明表达式 $p \vee q \vee r$ 是有定义的。因为先析取 p 和 q 再与 r 析取或先析取 q 和 r 再与 p 析取，其结果一样。同样，$p \wedge q \wedge r$ 也是有定义的。扩展这一推理过程可以证明：$p_1 \vee p_2 \vee \cdots \vee p_n$ 和 $p_1 \wedge p_2 \wedge \cdots \wedge p_n$ 均有定义，只要 p_1，p_2，…，p_n 为命题。进而可以注意到德摩根律可以扩展为

$$\neg(p_1 \vee p_2 \vee \cdots \vee p_n) \Leftrightarrow (\neg p_1 \wedge \neg p_2 \wedge \cdots \wedge \neg p_n)$$

和

$$\neg(p_1 \wedge p_2 \wedge \cdots \wedge p_n) \Leftrightarrow (\neg p_1 \vee \neg p_2 \vee \cdots \vee \neg p_n)$$

证明这些恒等式的方法将在第 3 章 3.3 节给出。

Extra Examples 表 1-13 中的逻辑等价关系以及业已建立起来的其他（如表 1-14 和表 1-15 中所示的那些）等价关系，可以用于构造其他等价关系，原因是复合命题中的一个命题可以用与

⊖ 这些恒等式是布尔代数中恒等式的一种特殊情况。请将这些恒等式与 1.7 节表 1-21 中的集合恒等式比较一下，也与 10.1 节表 10-5 中的布尔恒等式比较一下。

它逻辑等价的命题替换而不改变复合命题的真值。这种方法可由例 5 和例 6 得到说明。在这两个例子中，我们还使用了如下事实(见练习 50)：如果 p 和 q 逻辑等价，q 和 r 逻辑等价，那么 p 和 r 也逻辑等价。

表 1-13 逻辑等价

等价关系	名 称
$p \wedge \mathbf{T} \Leftrightarrow p$ $p \vee \mathbf{F} \Leftrightarrow p$	恒等律
$p \vee \mathbf{T} \Leftrightarrow \mathbf{T}$ $p \wedge \mathbf{F} \Leftrightarrow \mathbf{F}$	支配律
$p \vee p \Leftrightarrow p$ $p \wedge p \Leftrightarrow p$	幂等律
$\neg(\neg p) \Leftrightarrow p$	双非律
$p \vee q \Leftrightarrow q \vee p$ $p \wedge q \Leftrightarrow q \wedge p$	交换律
$(p \vee q) \vee r \Leftrightarrow p \vee (q \vee r)$ $(p \wedge q) \wedge r \Leftrightarrow p \wedge (q \wedge r)$	结合律
$p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$ $p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$	分配律
$\neg(p \wedge q) \Leftrightarrow \neg p \vee \neg q$ $\neg(p \vee q) \Leftrightarrow \neg p \wedge \neg q$	德摩根定律
$p \vee (p \wedge q) \Leftrightarrow p$ $p \wedge (p \vee q) \Leftrightarrow p$	吸收律
$p \vee \neg p \Leftrightarrow \mathbf{T}$ $p \wedge \neg p \Leftrightarrow \mathbf{F}$	否定律

表 1-14 涉及蕴含的逻辑等价

$p \to q \Leftrightarrow \neg p \vee q$
$p \to q \Leftrightarrow \neg q \to \neg p$
$p \vee q \Leftrightarrow \neg p \to q$
$p \wedge q \Leftrightarrow \neg(p \to \neg q)$
$\neg(p \to q) \Leftrightarrow p \wedge \neg q$
$(p \to q) \wedge (p \to r) \Leftrightarrow p \to (q \wedge r)$
$(p \to r) \wedge (q \to r) \Leftrightarrow (p \vee q) \to r$
$(p \to q) \vee (p \to r) \Leftrightarrow p \to (q \vee r)$
$(p \to r) \vee (q \to r) \Leftrightarrow (p \wedge q) \to r$

表 1-15 涉及双蕴含的逻辑等价

$p \leftrightarrow q \Leftrightarrow (p \to q) \wedge (q \to p)$
$p \leftrightarrow q \Leftrightarrow \neg p \leftrightarrow \neg q$
$p \leftrightarrow q \Leftrightarrow (p \wedge q) \vee (\neg p \wedge \neg q)$
$\neg(p \leftrightarrow q) \Leftrightarrow p \leftrightarrow \neg q$

例 5 证明 $\neg(p \vee (\neg p \wedge q))$ 和 $\neg p \wedge \neg q$ 逻辑等价。

解 我们可用真值表证明这两个复合命题等价。不过我们不这样做，而是每次使用表 1-13 中的一个等价关系，依次建立一串等价。从 $\neg(p \vee (\neg p \wedge q))$ 开始，到 $\neg(p \wedge \neg q)$ 结束，最终证明它们等价。我们有下列等价关系。

$$
\begin{aligned}
\neg(p \vee (\neg p \wedge q)) &\Leftrightarrow \neg p \wedge \neg(\neg p \wedge q) && \text{由第二德摩根定律} \\
&\Leftrightarrow \neg p \wedge [\neg(\neg p) \vee \neg q] && \text{由第一德摩根定律} \\
&\Leftrightarrow \neg p \wedge (p \vee \neg q) && \text{由双非律} \\
&\Leftrightarrow (\neg p \wedge p) \vee (\neg p \wedge \neg q) && \text{由分配律} \\
&\Leftrightarrow \mathbf{F} \vee (\neg p \wedge \neg q) && \text{由于} \neg p \wedge p \Leftrightarrow \mathbf{F} \\
&\Leftrightarrow (\neg p \wedge \neg q) \vee \mathbf{F} && \text{由析取的交换律} \\
&\Leftrightarrow \neg p \wedge \neg q && \text{由 F 的恒等律}
\end{aligned}
$$

于是 $\neg(p \vee (\neg p \wedge q))$ 和 $\neg p \wedge \neg q$ 逻辑等价。 ■

例 6 证明 $(p \wedge q) \to (p \vee q)$ 为永真式。

解 为证明这个命题是永真式，我们将用逻辑等价证明它逻辑上等价于 **T**。(注意：这也可以用真值表来完成。)

$$
\begin{aligned}
(p \wedge q) \to (p \vee q) &\Leftrightarrow \neg(p \wedge q) \vee (p \vee q) && \text{由例 3} \\
&\Leftrightarrow (\neg p \vee \neg q) \vee (p \vee q) && \text{由第一德摩根定律} \\
&\Leftrightarrow (\neg p \vee p) \vee (\neg q \vee q) && \text{由析取的结合律和交换律} \\
&\Leftrightarrow \mathbf{T} \vee \mathbf{T} && \text{由例 1 和析取的交换律} \\
&\Leftrightarrow \mathbf{T} && \text{由支配律}
\end{aligned}
$$

■

真值表可以用于判定复合命题是否为永真式。对于只含少数变量的命题，可以用手工完成这

一工作。但当变量数目增长时，就不可行了。例如，对于含 20 个变量的命题，它的真值表就有 $2^{20}=1\,048\,576$ 行。显然，你需要一台计算机帮助你以这种方式判定含 20 个变量的命题是否为永真式。但是当变量数为 1000 时，一台计算机能在一个可以接受的时间内判定复合命题是否为永真式吗？要检查 2^{1000} 种（这是一个超过 300 位的十进制数）可能的真值组合中的每一种，一台计算机在几万亿年之内都不可能完成。而且迄今尚没有其他已知的计算过程能使计算机在合理的时间之内判定变量数这么大的命题是否为永真式。在第 2 章我们学习算法复杂性时，将研究这一类问题。

练习

1. 用真值表证明下列等价关系。
   a) $p \wedge \mathbf{T} \Leftrightarrow p$　　b) $p \vee \mathbf{F} \Leftrightarrow p$　　c) $p \wedge \mathbf{F} \Leftrightarrow \mathbf{F}$
   d) $p \vee \mathbf{T} \Leftrightarrow \mathbf{T}$　　e) $p \vee p \Leftrightarrow p$　　f) $p \wedge p \Leftrightarrow p$
2. 证明 $\neg(\neg p)$ 和 p 逻辑等价。
3. 用真值表证明交换律。
   a) $p \vee q \Leftrightarrow q \vee p$　　b) $p \wedge q \Leftrightarrow q \wedge p$
4. 用真值表证明结合律
   a) $(p \vee q) \vee r \Leftrightarrow p \vee (q \vee r)$　　b) $(p \wedge q) \wedge r \Leftrightarrow p \wedge (q \wedge r)$
5. 用真值表证明分配律。
   $p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$
6. 用真值表证明等价关系。
   $\neg(p \wedge q) \Leftrightarrow \neg p \vee \neg q$
7. 用真值表证明下列各蕴含关系为永真式。
   a) $(p \wedge q) \rightarrow p$　　b) $p \rightarrow (p \vee q)$　　c) $\neg p \rightarrow (p \rightarrow q)$
   d) $(p \wedge q) \rightarrow (p \rightarrow q)$　　e) $\neg(p \rightarrow q) \rightarrow p$　　f) $\neg(p \rightarrow q) \rightarrow \neg q$
8. 用真值表证明下列蕴含关系为永真式。
   a) $[\neg p \wedge (p \vee q)] \rightarrow q$　　b) $[(p \rightarrow q) \wedge (q \rightarrow r)] \rightarrow (p \rightarrow r)$
   c) $[p \wedge (p \rightarrow q)] \rightarrow q$　　d) $[(p \vee q) \wedge (p \rightarrow r) \wedge (q \rightarrow r)] \rightarrow r$
9. 不用真值表证明练习 7 中的各蕴含关系为永真式。
10. 不用真值表证明练习 8 中的各蕴含关系为永真式。
11. 用真值表证明下列称为吸收律的等价关系。
    a) $p \vee (p \wedge q) \Leftrightarrow p$　　b) $p \wedge (p \vee q) \Leftrightarrow p$
12. 判断 $(\neg p \wedge (p \rightarrow q)) \rightarrow \neg q$ 是否为永真式。
13. 判断 $(\neg q \wedge (p \rightarrow q)) \rightarrow \neg p$ 是否为永真式。
14. 证明 $p \leftrightarrow q$ 和 $(p \wedge q) \vee (\neg p \wedge \neg q)$ 等价。
15. 证明 $(p \rightarrow q) \rightarrow r$ 和 $p \rightarrow (q \rightarrow r)$ 不等价。
16. 证明 $p \rightarrow q$ 和 $\neg q \rightarrow \neg p$ 逻辑等价。
17. 证明 $\neg p \leftrightarrow q$ 和 $p \leftrightarrow \neg q$ 逻辑等价。
18. 证明 $\neg(p \oplus q)$ 和 $p \leftrightarrow q$ 逻辑等价。
19. 证明 $\neg(p \leftrightarrow q)$ 和 $\neg p \leftrightarrow q$ 逻辑等价。
20. 证明 $(p \rightarrow q) \wedge (p \rightarrow r)$ 和 $p \rightarrow (q \wedge r)$ 逻辑等价。
21. 证明 $(p \rightarrow r) \wedge (q \rightarrow r)$ 和 $(p \vee q) \rightarrow r$ 逻辑等价。
22. 证明 $(p \rightarrow q) \vee (p \rightarrow r)$ 和 $p \rightarrow (q \vee r)$ 逻辑等价。
23. 证明 $(p \rightarrow r) \vee (q \rightarrow r)$ 和 $(p \wedge q) \rightarrow r$ 逻辑等价。
24. 证明 $\neg p \rightarrow (q \rightarrow r)$ 和 $q \rightarrow (p \vee r)$ 逻辑等价。
25. 证明 $p \leftrightarrow q$ 和 $(p \rightarrow q) \wedge (q \rightarrow p)$ 逻辑等价。
26. 证明 $p \leftrightarrow q$ 和 $\neg p \leftrightarrow \neg q$ 逻辑等价。
27. 证明 $\neg(p \leftrightarrow q)$ 和 $p \leftrightarrow \neg q$ 逻辑等价。

28. 证明$(p\vee q)\wedge(\neg p\vee r)\rightarrow(q\vee r)$是永真式。
29. 证明$(p\rightarrow q)\wedge(q\rightarrow r)\rightarrow(p\rightarrow r)$是永真式。

只含逻辑运算符$\vee$、$\wedge$和$\neg$的复合命题的对偶是将该命题中的每个$\vee$用$\wedge$代替，每个$\wedge$用$\vee$代替，每个**T**用**F**代替，每个**F**用**T**代替得到的命题。命题s的对偶用s^*表示。

30. 求下列命题的对偶。

    a) $p\wedge\neg q\wedge\neg r$　　b) $(p\wedge q\wedge r)\vee s$　　c) $(p\vee\mathbf{F})\wedge(q\vee\mathbf{T})$
31. 证明$(s^*)^*=s$
32. 表 1-13 中的逻辑等价关系除双非律外都是成对的。证明每一对命题都是互为对偶的。

**33. 为什么两个只含运算符$\wedge$、$\vee$和$\neg$的等价复合命题的对偶也等价？

34. 找一个只含命题p、q和r的复合命题，当p和q为真而r为假时命题为真，否则为假。[提示：用各个命题或其否定构造合取。]
35. 找一个只含命题p、q和r的复合命题，在p、q和r中恰有两个为真时命题为真，否则为假。[提示：构造合取的析取。对命题成真的每一种组合包含一个合取。每个合取都应包含三个命题中每个命题或它的否定。]
36. 假定用n个命题变量给出一个真值表。证明可依此表构造一个复合命题，使其真值与此表一致。构造办法是：对变量的每一种使复合命题成真的组合，用变量或变量的否定作成合取；再对所有合取作析取。这样得到的复合命题称为*析取范式*。

如果每一个复合命题都逻辑等价于一个只含某些运算符的复合命题，则这一组逻辑运算符称为*功能完备的*。

37. 证明$\neg$、$\wedge$和$\vee$构成一个功能完备的逻辑运算符集合。[提示：利用练习 26 中给出的事实，即每个命题都逻辑等价于一个析取范式。]

*38. 证明$\neg$和$\wedge$构成一个功能完备的逻辑运算符集合。[提示：首先用德摩根定律证明$p\vee q$等价于$\neg(\neg p\wedge\neg q)$。]

*39. 证明$\neg$和$\vee$构成一个功能完备的逻辑运算符集合。

下面几道题用到逻辑运算符 NAND(与非)和 NOR(或非)。命题p NAND q在p为假或q为假或两者均为假时为真；当p和q均为真时为假。命题p NOR q只在p和q均为假时为真，否则为假。命题p NAND q和p NOR q分别表示为$p\mid q$和$p\downarrow q$。(运算符$\mid$和$\downarrow$分别以谢佛和皮尔斯的名字命名为*谢佛竖*(Sheffer stroke)和*皮尔斯*[⊖]*箭头*。)

40. 为运算符 NAND 构造真值表。



⊖ **查理·皮尔斯(Charles Sanders Peirce, 1839—1914)** 许多人都认为查理·皮尔斯是美国最有创造性和最多才多艺的知识分子。他生于麻省的剑桥，其父本杰明·皮尔斯是哈佛大学的数学和自然哲学教授。皮尔斯就学于哈佛(1855—1859 年)并获文学硕士学位(1862)，后在劳伦斯科学学校获化学高级学位(1863)。他父亲鼓励他从事自然科学，但他却选择了研究逻辑和科学方法论。

1861 年，为更好理解科学方法论，皮尔斯当上了美国海岸观测署的助理。他在观测署的服务使他在南北战争期间免于从军。在该署工作期间，皮尔斯进行了天文和大地测量工作。他应用椭圆函数理论的最新数学成果，对钟摆设计和地图投影做出了奠基性的贡献。他是第一个把光的波长作为度量单位的人。皮尔斯被提升为观测署助理署长，并担任此职直到 1891 年被迫退职，因为他不同意观测署新的行政当局设定的工作方向。

尽管皮尔斯毕生致力于物理科学，他仍然提出了各种科学的一个层次结构，其中数学位于最高层，而且一门科学的方法可以被其下层科学采用。他还是美国实用主义哲学理论的奠基人。

皮尔斯唯一的学术职位是 1879～1884 年在巴尔的摩的约翰斯·霍普金斯大学担任逻辑学讲师。这期间他完成的数学工作包括他对逻辑、集合论、抽象代数和数学原理的贡献。他的工作至今仍产生影响，他在逻辑上的某些工作近来已被应用于人工智能。皮尔斯相信，研究数学可以开发大脑的想象力、抽象思维能力和归纳能力。他从观测署退休以后的五花八门的工作包括为报纸和期刊写稿、编著多部学术辞典、翻译科技论文、客座授课及撰写教科书。不幸的是，以这些工作得到的收入不足以使他和他的第二任妻子免受贫穷之苦。晚年他得到由他的许多赞赏者创立的基金的支持，该基金由他终生的朋友哲学家威廉·詹姆斯(William James)管理。尽管皮尔斯就广泛的主题写作并出版了大量著作，他仍然留下了 100 000 多页未出版的手稿。由于这些未发表的作品很难读，学者们只是在近来才开始理解他大量贡献中的一部分。一些人正献身于把他的工作推上因特网，从而使皮尔斯对全世界的贡献更好地得到鉴赏。

41. 证明 $p \mid q$ 逻辑等价于 $\neg(p \wedge q)$。
42. 为运算符 *NOR* 构造真值表。
43. 证明 $p \downarrow q$ 逻辑等价于 $\neg(p \vee q)$。
44. 本题将证明 $\{\downarrow\}$ 是功能完备的一个逻辑运算符集合。
    a)证明 $p \downarrow p$ 逻辑等价于 $\neg p$。
    b)证明 $(p \downarrow q) \downarrow (p \downarrow q)$ 逻辑等价于 $p \vee q$。
    c)从练习 29 和本题的 a)、b)得知 $\{\downarrow\}$ 是功能完备的一个逻辑运算符集合。

*45. 只用运算符 $\downarrow$ 构造一个等价于 $p \to q$ 的命题。
46. 证明 $\{\mid\}$ 是功能完备的一个逻辑运算符集合。
47. 证明 $p \mid q$ 和 $q \mid p$ 等价。
48. 证明 $p \mid (q \mid r)$ 和 $(p \mid q) \mid r$ 不等价。（因此，逻辑运算符 $\mid$ 不满足结合律。）

*49. 只用命题 p 和 q 能有多少不同的复合命题真值表？
50. 如果复合命题 p、q 和 r 中，p 与 q 逻辑等价，q 与 r 逻辑等价，证明 p 与 r 逻辑等价。
51. 下面的语句取自一个电话系统的规范说明："如果电话号码数据库是打开的，那么监督程序被置于关闭状态，只要系统不在初态。"这句话有两个蕴含，使规范说明很难懂。找一个等价的易懂的规范说明，使其只涉及析取和否定，不涉及蕴含。
52. 通过对 p、q、r 赋真值，析取表达式 $p \vee \neg q$，$\neg p \vee q$，$q \vee r$，$q \vee \neg r$，$\neg q \vee \neg r$ 中有多少可以同时为真？
53. 通过对 p、q、r、s 赋真值，析取表达式 $p \vee \neg q \vee s$，$\neg p \vee \neg r \vee s$，$\neg p \vee \neg r \vee \neg s$，$\neg p \vee q \vee \neg s$，$q \vee r \vee \neg s$，$q \vee \neg r \vee \neg s$，$\neg p \vee \neg q \vee \neg s$，$p \vee r \vee s$，$p \vee r \vee \neg s$ 中有多少可以同时为真？

    如果对命题变量的真值赋值可使复合命题为真，则称这个复合命题是*可满足的*。

54. 下述复合命题哪些是可满足的？
    a) $(p \vee q \vee \neg r) \wedge (p \vee \neg q \vee \neg s) \wedge (p \vee \neg r \vee \neg s) \wedge (\neg p \vee \neg q \vee \neg s) \wedge (p \vee q \vee \neg s)$
    b) $(\neg p \vee \neg q \vee r) \wedge (\neg p \vee q \vee \neg s) \wedge (p \vee \neg q \vee \neg s) \wedge (\neg p \vee \neg r \vee \neg s) \wedge (p \vee q \vee \neg r) \wedge (p \vee \neg r \vee \neg s)$
    c) $(p \vee q \vee r) \wedge (p \vee \neg q \vee \neg s) \wedge (q \vee \neg r \vee s) \wedge (\neg p \vee r \vee s) \wedge (\neg p \vee q \vee \neg s) \wedge (p \vee \neg q \vee \neg r) \wedge (\neg p \vee \neg q \vee s) \wedge (\neg p \vee \neg r \vee \neg s)$
55. 解释一下用于判定复合命题可满足性的算法如何能用来判断一个复合命题是否是永真式。（提示：考虑 $\neg p$，其中 p 是已经考察过的命题。）

1.3 谓词和量词

1.3.1 引言

含变量的语句，如

$$\text{“}x > 3\text{”，“}x = y + 3\text{” 和 “}x + y = z\text{”}$$

常见于数学断言和计算机程序。在变量值未知的时候，这些语句既不成真也不为假。本节将讨论从这种语句产生命题的方式。

语句"x 大于 3"有两部分，第一部分即变量 x 是语句的主语，第二部分*谓词*"大于 3"表明语句的主语会有的一个性质。我们可以用 $P(x)$ 表示语句"x 大于 3"，其中 P 表示谓词"大于 3"，而 x 是变量。也把语句 $P(x)$ 说成是*命题函数* P 在 x 的值。一旦给变量 x 赋一个值，语句 $P(x)$ 就成为命题，因而有真值。考虑下面的例子。

例 1 令 $P(x)$ 表示语句"$x>3$"，$P(4)$ 和 $P(2)$ 的真值是什么？

解 在语句"$x>3$"中让 $x=4$ 即得到语句 $P(4)$。因此 $P(4)$ 即语句"$4>3$"，为真；但语句

$p(2)$即"$2>3$"，为假。■

还可以让语句中含不止一个变量，例如，考虑语句"$x=y+3$"。我们可以用$Q(x, y)$表示这个语句，其中x，y为变量，Q为谓词。在赋值给x和y时，语句$Q(x, y)$就有了真值。

例2 令$Q(x, y)$表示语句"$x=y+3$"，命题$Q(1, 2)$和$Q(3, 0)$的真值是什么？

解 要得到$Q(1, 2)$，在$Q(x, y)$中令$x=1$，$y=2$。因此，$Q(1, 2)$为语句"$1=2+3$"，为假。语句$Q(3, 0)$为语句"$3=0+3$"，为真。■

同样，可以用$R(x, y, z)$表示语句"$x+y=z$"，在x、y和z被赋值时，这一语句就有了真值。

例3 命题$R(1, 2, 3)$和$R(0, 0, 1)$的真值是什么？

解 在语句$R(x, y, z)$中令$x=1$，$y=2$，$z=3$，即得到命题$R(1, 2, 3)$。可以看出$R(1, 2, 3)$就是语句"$1+2=3$"，为真。$R(0, 0, 1)$即语句"$0+0=1$"，为假。■

一般涉及n个变量如$x_1, x_2, \cdots, x_n$的语句可以用$P(x_1, x_2, \cdots, x_n)$表示。

形为$P(x_1, x_2, \cdots, x_n)$的语句是**命题函数**P在n元组$(x_1, x_2, \cdots, x_n)$的值，P也称为**谓词**。

下面的例子说明，命题函数可以出现在计算机程序中。

例4 考虑语句

$$\textbf{if } x>0 \textbf{ then } x := x+1$$

当程序中遇到这样一条语句时，变量x在程序运行到此刻的值即被代入$P(x)$，也就是代入"$x>0$"。如果对x的这一值$P(x)$为真，即执行赋值语句$x:=x+1$，于是x的值增加1。如果对x的这一值$P(x)$为假，则不执行赋值语句，所以x的值不改变。■

1.3.2 量词

当命题函数中所有变量均被赋值时，得到的命题有一个真值。还有另一重要方式，也可以从命题函数产生命题，这就是**量化**。本节将讨论两类量化，即全称量化和存在量化。与谓词和量词相关的逻辑领域称为**谓词演算**。

全称量词 许多数学语句断定某一性质对变量在某一特定域内的所有值均为真，这一特定域称为变量的**论域**。这类语句用全称量化表示。命题函数的全称量化是这样一个命题：它断言$P(x)$对x在其论域中的所有值为真，论域规定变量x可能取的值。

定义1 $P(x)$的**全称量化**是命题"$P(x)$对x在其论域的所有值为真"。

符号

$$\forall x P(x)$$

表示$P(x)$的全称量化，其中$\forall$称为**全称量词**。命题$\forall x P(x)$也可表示为"对所有x，$P(x)$"或"对每个x，$P(x)$"。

> **注意** 最好不用"任何"一词，因为它常引起歧义，不知指的是"每一个"还是"某一个"。在有些情况下，"任何"没有二义性，例如，用于否定句时，像"没有任何理由不努力学习。"

下面几个例子说明了全称量词的使用。

例5 令$P(x)$为语句"$x+1>x$"，量化语句$\forall x P(x)$的真值是什么？其中论域是实数

集合。

解 由于 $P(x)$ 对所有实数 x 均为真，量化语句

$$\forall xP(x)$$

的值为真。 ■

例 6 令 $Q(x)$ 表示语句“$x<2$”。若论域是实数集合，量化语句 $\forall xQ(x)$ 的真值是什么？

解 $Q(x)$ 并非对所有实数都为真，例如，$Q(3)$ 就为假，因此

$$\forall xQ(x)$$

为假。 ■

当论域中的所有元素可以一一列出——即例如列成 x_1，x_2，…，x_n——时，量化语句 $\forall xP(x)$ 与合取

$$P(x_1) \wedge P(x_2) \wedge \cdots \wedge P(x_n)$$

是一回事，因为这一合取为真当且仅当 $P(x_1)$，$P(x_2)$，…，$P(x_n)$ 全为真。

例 7 若论域是不超过 4 的正整数，$P(x)$ 是语句“$x^2<10$”，$\forall xP(x)$ 的真值是什么？

解 语句 $\forall xP(x)$ 就是合取

$$P(1) \wedge P(2) \wedge P(3) \wedge P(4)$$

因为论域由 1、2、3 和 4 组成。由于 $P(4)$ 即语句“$4^2<10$”为假，$\forall xP(x)$ 为假。 ■

例 8 若 $T(x)$ 是“x 有双亲”且论域包括所有人，则语句 $\forall xT(x)$ 是什么含义？

解 语句 $\forall xT(x)$ 的意思是对于每一个人 x，这个人有双亲。用自然语言表达就是“每个人都有双亲”。这个语句是真的(除了克隆人)。 ■

当使用量词时，说明论域是很重要的。量化语句的真值通常取决于哪个元素处于论域中，如例 9 所示。

例 9 如果论域包括所有实数，$\forall x(x^2 \geqslant x)$ 的真值是什么？如果论域包括所有整数，真值又是什么？

解 注意 $x^2 \geqslant x$ 当且仅当 $x^2-x=x(x-1) \geqslant 0$。因此，$x^2 \geqslant x$ 当且仅当 $x \leqslant 0$ 或 $x \geqslant 1$。由此得出，如果论域包括所有实数，$\forall x(x^2 \geqslant x)$ 为假(因为当 $0<x<1$ 时，不等式不成立)。然而，如果论域为整数，$\forall x(x^2 \geqslant x)$ 为真，因为对任一整数 x 都不存在 $0<x<1$。 ■

要证明形如 $\forall xP(x)$ 的语句为假，其中 $P(x)$ 是一个命题函数，我们只需要在论域中找到 x 的一个值，此时 $P(x)$ 为假。x 的这个值称为语句 $\forall xP(x)$ 的*反例*。

例 10 假设 $P(x)$ 是“$x^2>0$”。要证明语句 $\forall xP(x)$ 为假(其中论域是所有整数)，我们给出一个反例。$x=0$ 是一个反例，因为当 $x=0$ 时，$x^2=0$ 而不大于 0。 ■

在数学研究中，寻找全称量化语句的反例是一个重要的部分，在本书后续章节中还会介绍。

存在量词 许多数学语句断定有一个具有某种性质的元素。这类语句用存在量化表示。用存在量化可以构成这样一个命题：它成真的充分必要条件是论域中至少有一个值使 $P(x)$ 为真。

定义 2 $P(x)$ 的*存在量化*是命题“论域中存在一个元素 x 使 $P(x)$ 为真”。

我们用符号

$$\exists xP(x)$$

表示 $P(x)$ 的存在量化，其中 $\exists$ 称为*存在量词*。存在量化 $\exists xP(x)$ 也可表示为

“有一个 x 使得 $P(x)$，”

"至少有一个 x 使得 $P(x)$，"

或

"对某个 x，$P(x)$。"

下面几个例子说明了存在量词的使用。

Extra Examples **例 11** 令 $P(x)$ 表示语句"$x>3$"，论域为实数集合，量化语句 $\exists xP(x)$ 的真值是什么？

解 因为"$x>3$"在如 $x=4$ 时为真，$P(x)$ 的存在量化 $\exists xP(x)$ 为真。 ■

例 12 令 $Q(x)$ 表示语句"$x=x+1$"，论域是实数集，量化语句 $\exists xQ(x)$ 的真值是什么？

解 因为对每个实数 x，$Q(x)$ 都为假，$Q(x)$ 的存在量化 $\exists xQ(x)$ 为假。 ■

当论域中的所有元素可以一一列出——例如列成 x_1，x_2，…，x_n——时，存在量化 $\exists xP(x)$ 与析取

$$P(x_1) \vee P(x_2) \vee \cdots \vee P(x_n)$$

是一回事，因为这一析取当且仅当 $P(x_1)$，$P(x_2)$，…，$P(x_n)$ 中至少一个为真时为真。

例 13 若 $P(x)$ 是语句"$x^2>10$"，论域为不超过 4 的正整数，$\exists xP(x)$ 的真值是什么？

解 由于论域为{1，2，3，4}，命题 $\exists xP(x)$ 就是析取

$$P(1) \vee P(2) \vee P(3) \vee P(4)$$

由于 $P(4)$ 即"$4^2>10$"为真，$\exists xP(x)$ 为真。 ■

表 1-16 总结了全称量词和存在量词的含义。

表 1-16 量词

语 句	何时为真？	何时为假？
$\forall xP(x)$	对每一个 x，$P(x)$ 都为真	有一个 x，使 $P(x)$ 为假
$\exists xP(x)$	有一个 x，使 $P(x)$ 为真	对每一个 x，$P(x)$ 都为假

在决定量化语句的真值时，借助循环与搜索来思考是有益的。假定在变量 x 的论域中有 n 个对象，要决定 $\forall xP(x)$ 是否为真，可以对 x 的 n 个值循环查看 $P(x)$ 是否总是真。如果遇到 x 的一个值使 $P(x)$ 为假，那就证明 $\forall xP(x)$ 为假，否则 $\forall xP(x)$ 为真。要决定 $\exists xP(x)$ 是否为真，我们循环查看 x 的 n 个值，搜索使 $P(x)$ 成真的 x 之值。如果找到一个，那么 $\exists xP(x)$ 为真；如果总也找不到这样的 x，则判定 $\exists xP(x)$ 为假。（注意，当论域有无穷多个值时，这一搜索过程不适用。不过以这种方式思考量化语句的真值仍是有益的。）

1.3.3 绑定变量

当量词作用于变量 x 或给这一变量赋值时，我们说此变量的这一次出现为*绑定的*。没有被量词绑定或设置为与某一特定值相等的变量出现称为*自由的*。出现在命题函数中的所有变量必须是绑定的，才能把此命题函数转变为命题。可以用全称量词、存在量词和赋值来完成转变。

逻辑表达式中应用量词的部分称为这个量词的*作用域*。因此，一个变量是自由的——如果在指定这个变量的公式中，变量在所有量词的作用域之外。

例 14 在语句 $\exists xQ(x, y)$ 中，变量 x 通过存在量词 $\exists x$ 被绑定，但变量 y 是自由的，因为

它没有被量词绑定，且没有对它进行赋值。

在语句 $\exists x(P(x) \land Q(x)) \lor \forall xR(x)$ 中，所有变量都绑定了。第一个量词 $\exists x$ 的作用域是表达式 $P(x) \land Q(x)$，因为 $\exists x$ 只作用于 $P(x) \land Q(x)$。类似地，第二个量词 $\forall x$ 的作用域是表达式 $R(x)$。也就是说，存在量词在 $P(x) \land Q(x)$ 中绑定变量 x，全称量词 $\forall x$ 在 $R(x)$ 中绑定变量 x。由此，我们可以用两个不同的变量 x 和 y 将语句写为 $\exists x(P(x) \land Q(x)) \lor \forall yR(y)$，因为两个量词的作用域不重叠。读者在日常使用中应注意这一点，相同的字母经常被用于表示被不同量词(其作用域不重叠)绑定的变量。 ■

1.3.4 否定

我们常会考虑量化表达式的否定，例如，考虑语句

“班上每个学生都学过一门微积分课”

的否定。这个语句是全称量化语句，即

$$\forall xP(x)$$

其中 $P(x)$ 为语句“x 学过一门微积分课”。这一语句的否定是“并非班上每个学生都学过一门微积分课”。这等价于“班上有个学生没有学过微积分课”。这也就是原命题函数之否定的存在量化，即

$$\exists x \neg P(x)$$

这个例子说明了下面的等价关系：

$$\neg \forall xP(x) \Leftrightarrow \exists x \neg P(x)$$

Extra Examples 假定我们想否定一个存在量化语句。例如，考虑命题“班上有个学生学过一门微积分课。”就是存在量化语句

$$\exists xQ(x)$$

其中 $Q(x)$ 为语句“x 学过一门微积分课”。这句话的否定是命题“并非班上有个学生学过微积分课”。这等价于“班上每个学生都没学过微积分课”，这也就是原命题函数之否定的全称量化，或用量词语言表示为

$$\forall x \neg Q(x)$$

这个例子说明了等价关系

$$\neg \exists xQ(x) \Leftrightarrow \forall x \neg Q(x)$$

表 1-17 给出了量词之否定的总结。

表 1-17 量词的否定

否　定	等价语句	何时为真?	何时为假?
$\neg \exists xP(x)$	$\forall x \neg P(x)$	对每个 x，$P(x)$为假	有 x，使 $P(x)$为真
$\neg \forall xP(x)$	$\exists x \neg P(x)$	有 x 使 $P(x)$为假	对每个 x，$P(x)$为真

注意 当谓词 $P(x)$的论域包含 n 个元素时，其中 n 是正整数，用于否定量化语句的规则恰与 1.2 节讨论的德摩根律相同。这是因为 $\neg \forall xP(x)$ 与 $\neg(P(x_1) \land P(x_2) \land \cdots \land P(x_n))$ 相等，而后者等价于 $\neg P(x_1) \lor \neg P(x_2) \lor \cdots \lor \neg P(x_n)$（由德摩根律），该式又等价于 $\exists x \neg P(x)$。类似地，$\neg \exists xP(x)$ 与 $\neg(P(x_1) \lor P(x_2) \lor \cdots \lor P(x_n))$ 相等，后者由德

摩根律等价于$\neg P(x_1) \wedge \neg P(x_2) \wedge \cdots \wedge \neg P(x_n)$，该式又等于$\forall x \neg P(x)$。

下面几个例子演示了量化语句的否定。

例 15 语句“有诚实的政治家”和“所有美国人都吃干酪汉堡包”的否定是什么？

解 令$H(x)$表示“x是诚实的”，则语句“有诚实的政治家”可以用$\exists x H(x)$来表示，其中论域包括所有政治家。这个语句的否定是$\neg \exists x H(x)$，它等价于$\forall x \neg H(x)$。这个否定可以表达为“每个政治家都不诚实。”

令$C(x)$为“x吃干酪汉堡包”，则语句“所有美国人都吃干酪汉堡包”可以用$\forall x C(x)$来表示，其中论域包括所有美国人。这个语句的否定是$\neg \forall x C(x)$，它等价于$\exists x \neg C(x)$。这个否定可以用几种不同的方式来表达，包括“一些美国人不吃干酪汉堡包”和“有一个美国人不吃干酪汉堡包”。 ■

例 16 语句$\forall x(x^2 > x)$和$\exists x(x^2 = 2)$的否定是什么？

解 $\forall x(x^2 > x)$的否定是语句$\neg \forall x(x^2 > x)$，它等价于$\exists x \neg (x^2 > x)$。这个表达式可以重写为$\exists x(x^2 \leqslant x)$。$\exists x(x^2 = 2)$的否定是语句$\neg \exists x(x^2 = 2)$，它等价于$\forall x \neg (x^2 = 2)$。这个表达式可以重写为$\forall x(x^2 \neq 2)$。这些语句的真值取决于论域。 ■

1.3.5 翻译语句为逻辑表达式

将汉语(或其他自然语言)语句翻译成逻辑表达式，这在数学、逻辑编程、人工智能、软件工程及许多其他学科中是一项重要的任务。在1.1节中，我们学习了把句子译成由命题和逻辑联接词构成逻辑表达式的过程。在其时，我们特意对要求谓词和量词的语句翻译避而不谈。当用到量词时，将语句翻译成逻辑表达式会变得更复杂。而且，有多种方式可用于翻译一个特定的语句。(因此，没有“菜谱”式的方法可供你按部就班地学习。)我们会给出一些例子说明如何将汉语语句翻译成逻辑表达式。翻译的目标是生成简单而有用的逻辑表达式。本节我们只讨论这样的语句，可将其翻译成使用单个量词的逻辑表达式。下一节会讨论一些更复杂的要求多个量词的语句。

例 17 使用谓词和量词表达语句“班上的每个学生都学过微积分课”。

解 首先，重写语句以确定要使用的量词。重写后可得：

“对班上的每一个学生，该学生学过微积分课。”

接着，引入变量x，语句就变成

“对班上的每一个学生x，x学过微积分课。”

然后，引入谓词$C(x)$，表示语句“x学过微积分课”。因此，如果x的论域是班上的学生，我们可以将语句翻译为$\forall x C(x)$。

然而，还有其他正确的翻译方法，并可使用不同的论域和其他谓词。具体选择什么方法取决于后续要进行的推理。例如，我们可能对更广泛的人群感兴趣，而不仅仅是班上的学生。如果将论域改成所有人，则语句就要变成

“对每个人x，如果x是班上的学生，那么x学过微积分课。”

如果$S(x)$表示语句x在这个班上，则我们的语句可表达为$\forall x(S(x) \rightarrow C(x))$。(注意：语句不能被表达为$\forall x(S(x) \wedge C(x))$，因为这表示的是所有人都是这个班上的学生并且学过微积分课。)

最后，如果我们对学生除微积分之外的其他主修课程感兴趣的话，可以使用双变量量词

$Q(x, y)$表示语句“学生 x 学过课程 y”。这样在上述两种方法中我们就要把 $C(x)$替换成 $Q(x$，微积分$)$，得到 $\forall xQ(x$，微积分$)$或 $\forall x(S(x) \to Q(x$，微积分$))$。■

在例 17 中我们演示了采用谓词和量词表达相同语句的不同方法。不过要记住，应该总是采用适合后续推理使用的最简单的方法。

例 18 用谓词和量词表达语句“这个班上某个学生去过墨西哥”和“这个班上每个学生或去过加拿大，或去过墨西哥。”

解 引入谓词 $M(x)$表示语句“x 去过墨西哥”。如果 x 的论域是这个班上的学生，我们就可以将第一个语句翻译为 $\exists xM(x)$。

然而，如果 x 的论域是所有人，我们引入谓词 $S(x)$表示语句“x 是这个班上的一个学生”。答案就变成了 $\exists x(S(x) \land M(x))$，因为它表示有某个人 x 是这个班上的一个学生并且去过墨西哥。

类似地，对第二个语句，我们令 $C(x)$表示语句“x 去过加拿大”。由前面的推理，如果 x 的论域是这个班上的学生，则第二个语句可以表达为 $\forall x(C(x) \lor M(x))$。然而，如果 x 的论域是所有人，答案就应该是 $\forall x(S(x) \to (C(x) \lor M(x)))$。

不同于使用谓词 $M(x)$和 $C(x)$分别表示 x 去过墨西哥和 x 去过加拿大，我们可以使用两个变量谓词 $V(x, y)$表示“x 去过国家 y”。在本例中，$V(x$，墨西哥$)$和 $V(x$，加拿大$)$与 $M(x)$和 $C(x)$意思相同并可以代替它们。如果碰到某人去过多个国家的语句，用这种双变量的方法更可取。否则为了简单将坚持用单变量谓词 $M(x)$和 $C(x)$。■

1.3.6 选自 Lewis Carroll 的例子

Lewis Carroll（实际上是 C. L. Dodgson[⊖] 的笔名）是《*Alice in Wonderland*》的作者，也是有关符号逻辑几篇论文的作者。他的书中有用量词推理的例子。下面的两个例子选自他的书《*Symbolic logic*》，选自这本书的其他例子在本节末的练习中给出。这些例子说明了怎样用量词表示各种类型的词句。

例 19 考虑下面这些语句，其中头两句称为前提，第三句称为结论。作为一个整体它们被称为一个论证。

“所有狮子都是凶猛的。”

“有些狮子不喝咖啡。”

“有些凶猛的动物不喝咖啡。”

（1.5 节将讨论判定结论部分是不是前提部分有效后果的问题，就这个例子来说，结论是有效的。）令 $P(x)$，$Q(x)$和 $R(x)$分别为语句“x 是狮子”，“x 是凶猛的”和“x 喝咖啡”。假定所有动物的集合为论域，用量词及 $P(x)$、$Q(x)$和 $R(x)$表示上面这些语句。

解 可以将这些句子表示为：

$$\forall x(P(x) \to Q(x))$$

⊖ **道奇森（Charles Lutwidge Dodgson，1832—1898）** 我们知道 Charles Dodgson 即是 Lewis Carroll，后者是他发表逻辑著作的笔名。道奇森是牧师的儿子，在 11 个子女中排行第三。这 11 个全是结巴。在成年人的包围下他常感到不安。道奇森十分严肃和忠于宗教信仰。他与 Dean Liddell 的三个女儿的友谊使他写成《爱丽丝漫游奇境记》，这本书为他赢得了金钱和名声。

道奇森 1854 年毕业于牛津，并于 1857 年获得艺术硕士学位。1855 年他被任命为牛津 Christ Church 学院的数学讲师。1861 年英国教会委任他以圣职，但他从来没有实际履行过宗教职责。他的著作包括有关几何、行列式以及竞赛和选举中的数学问题的论文和书籍。（他还以笔名 Lewis Carroll 写过许多关于消遣性逻辑的作品。）

$$\exists x(P(x) \land \neg R(x))$$
$$\exists x(Q(x) \land \neg R(x))$$

注意，第二句不能表示为$\exists x(P(x) \rightarrow \neg R(x))$。原因是$P(x) \rightarrow \neg R(x)$在$x$不是狮子时总是成真，所以只要有一只不是狮子的动物，$\exists x(P(x) \rightarrow \neg R(x))$就成真，即使所有狮子都喝咖啡它也成真。同样，第三句也不能写成

$$\exists x(Q(x) \rightarrow \neg R(x))$$

■

例 20 考虑下面的语句，其中前 3 个语句为前提，第 4 个语句为有效结论。

“所有蜂鸟都五彩斑斓。”

“没有大鸟以蜜为生。”

“不以蜜为生的鸟都色彩单调。”

“蜂鸟都是小鸟。”

令$P(x)$，$Q(x)$，$R(x)$和$S(x)$分别为语句“x是只蜂鸟”，“x是大的”，“x以蜜为生”和“x五彩斑斓”。假定以所有鸟的集合为论域，用量词及$P(x)$、$Q(x)$、$R(x)$和$S(x)$表示上述语句。

解 可以把论证中的语句表示为

$$\forall x(P(x) \rightarrow S(x))$$
$$\neg \exists x(Q(x) \land R(x))$$
$$\forall x(\neg R(x) \rightarrow \neg S(x))$$
$$\forall x(P(x) \rightarrow \neg Q(x))$$

(注意，我们假定了“小”就是“不大”，“色彩单调”就是“不五彩斑斓”。为证明第 4 句是前 3 句的有效结论，需要用到将在 1.5 节讨论的推理规则。) ■

1.3.7 逻辑程序设计

有一类重要的程序设计语言，使用谓词逻辑的规则进行推理。其中之一就是 Prolog (Programming in Logic 的缩写)，这种语言开发于 20 世纪 70 年代，应用于人工智能领域。Prolog 程序包括一组声明，其中包括两类语句：Prolgo 事实和 Prolog 规则。Prolog 事实定义谓词，这是通过指定满足这些谓词的元素来实现的。Prolog 规则使用那些 Prolog 事实定义好的谓词来定义新的谓词。例 21 演示这些概念。

例 21 考虑一个 Prolog 程序，给出的事实是每门课程的教师和学生注册的课程。程序使用这些事实回答与教师授课情况相关的查询。这样的程序可使用谓词 instructor(p, c)和 enrolled(s, c)分别表示教授p是讲授课程c的老师及学生s注册课程c。例如，此程序中的 Prolog 事实可能包含：

```
instructor(chan, math273)
instructor(patel, ee222)
instructor(grossman, cs301)
enrolled(kevin, math273)
enrolled(juana, ee222)
enrolled(juana, cs301)
enrolled(kiko, math273)
enrolled(kiko, cs301)
```

(用小写字母表示项目是因为 Prolog 把以大写字母开始的名字当做变量。)

一个新的谓词 teaches(p, s)表示教授p教学生s，可以用 Prolog 规则来定义：

```
teaches(P, S):-instructor(P, C), enrolled(S, C)
```

上述语句意味着如果教授 p 是课程 c 的老师，学生 s 注册课程 c，teaches(p，s)为真。(注意，在 Prolog 中逗号用于表示谓词合取。类似地，分号表示谓词析取。)

Prolog 使用给定的事实和规则回答查询。例如，使用上述的事实和规则，查询

```
? enrolled(kevin, math273)
```

生成应答

```
yes
```

因为事实 enrolled(kevin，math273)是一个输入。查询

```
? enrolled(X, math273)
```

生成应答

```
kevin
kiko
```

要生成上面的应答，Prolog 确定 enrolled(X，math273)包含的作为 Prolog 事实的 X 所有可能的值。类似地，要找到 Juana 所选课程的教师，可以用查询

```
? teaches(X, juana)
```

这个查询返回

```
patel
grossman
```

■

练习

1. 令 $P(x)$ 表示语句"$x \leq 4$"。下列各项的真值是什么？
   a) $P(0)$　　b) $P(4)$　　c) $P(6)$
2. 令 $P(x)$ 表示语句"单词 x 含字母 a。"下列各项的真值是什么？
   a) P(orange)　　b) P(lemon)　　c) P(true)　　d) P(false)
3. 令 $Q(x, y)$ 表示语句"x 是 y 的首府。"下列各项的真值是什么？
   a) Q(丹佛，科罗拉多)　　b) Q(底特律，密歇根)　　c) Q(马萨诸塞，波士顿)　　d) Q(纽约，纽约)
4. 给出执行语句 if $P(x)$ then $x := 1$ 以后 x 的值，其中 $P(x)$ 为语句"$x>1$"，且执行到上述语句时 x 的值是：
   a) $x=0$　　b) $x=1$　　c) $x=2$
5. 令 $P(x)$ 为语句"x 每个工作日都花五个多小时上课"，其中 x 的论域是学生集合。用句子表达下列各量化语句。
   a) $\exists x P(x)$　　b) $\forall x P(x)$　　c) $\exists x \neg P(x)$　　d) $\forall x \neg P(x)$
6. 令 $N(x)$ 为语句"x 已经去过北达科他"，论域包括你学校的所有学生。用汉语表达下列各量词。
   a) $\exists x N(x)$　　b) $\forall x N(x)$　　c) $\neg \exists x N(x)$
   d) $\exists x \neg N(x)$　　e) $\neg \forall x N(x)$　　f) $\forall x \neg N(x)$
7. 将下列语句翻译成汉语语句，其中 $C(x)$ 是"x 是一个喜剧演员"，$F(x)$ 是"x 很有趣"，论域是所有人。
   a) $\forall x(C(x) \rightarrow F(x))$　　b) $\forall x(C(x) \wedge F(x))$　　c) $\exists x(C(x) \rightarrow F(x))$　　d) $\exists x(C(x) \wedge F(x))$

8. 将下列语句翻译成汉语语句，其中 $R(x)$是“x 是一只兔子”，$H(x)$是“x 跳跃”，论域是所有动物。

a) $\forall x(R(x) \rightarrow H(x))$　b) $\forall x(R(x) \wedge H(x))$　c) $\exists x(R(x) \rightarrow H(x))$　d) $\exists x(R(x) \wedge H(x))$

9. 令 $P(x)$为语句“x 说俄语”，$Q(x)$为语句，“x 了解计算机语言 C++”。用 $P(x)$、$Q(x)$、量词和逻辑联接词表示下列各句子，量词的论域为你校全体学生的集合。

a)你校有个学生既会说俄语又了解 C++。　b)你校有个学生会说俄语，但不了解 C++。

c)你校所有学生或会说俄语，或了解 C++。　d)你校没有学生会说俄语或了解 C++。

10. 令 $C(x)$为语句“x 有一只猫”，$D(x)$为语句“x 有一只狗”，$F(x)$为语句“x 有一只雪貂”。用$C(x)$、$D(x)$、$F(x)$、量词和逻辑联接词表达下述语句。令论域包括你班上的所有学生。

a)班上的一个学生有一只猫、一只狗和一只雪貂。

b)班上的所有学生有一只猫、一只狗或一只雪貂。

c)班上的一些学生有一只猫和一只雪貂，但没有狗。

d)班上没有学生同时有一只猫、一只狗和一只雪貂。

e)对猫、狗和雪貂这三种动物的任意一种，班上都有学生将其作为宠物。

11. 令 $P(x)$为语句“$x=x^2$”。如果论域是整数集合，下列各项的真值是什么？

a)$P(0)$　b)$P(1)$　c)$P(2)$

d)$P(-1)$　e) $\exists x P(x)$　f) $\forall x P(x)$

12. 令 $Q(x)$为语句“$x+1>2x$”。如果论域为整数集合，下列各项的真值是什么？

a)$Q(0)$　b)$Q(-1)$　c)$Q(1)$　d) $\exists x Q(x)$

e) $\forall x Q(x)$　f) $\exists x \neg Q(x)$　g) $\forall x \neg Q(x)$

13. 如果论域为整数集合，下列各语句的真值是什么？

a) $\forall n(n+1>n)$　b) $\exists n(2n=3n)$　c) $\exists n(n=-n)$　d) $\forall n(n^2 \geqslant n)$

14. 如果论域为实数集合，下列各语句的真值是什么？

a) $\exists x(x^3=-1)$　b) $\exists x(x^4<x^2)$　c) $\forall x((-x)^2=x^2)$　d) $\forall x(2x>x)$

15. 如果所有变量的论域为整数集合，下列各语句的真值是什么？

a) $\forall n(n^2 \geqslant 0)$　b) $\exists n(n^2=2)$　c) $\forall n(n^2 \geqslant n)$　b) $\exists n(n^2<0)$

16. 如果每个变量的论域都为实数集合，下列各语句的真值是什么？

a) $\exists x(x^2=2)$　b) $\exists x(x^2=-1)$　c) $\forall x(x^2+2 \geqslant 1)$　d) $\forall x(x^2 \neq x)$

17. 假设命题函数 $P(x)$的论域为整数 0，1，2，3 和 4。使用析取、合取和否定写出下列命题。

a) $\exists x P(x)$　b) $\forall x P(x)$　c) $\exists x \neg P(x)$

d) $\forall x \neg P(x)$　e)$\neg \exists x P(x)$　f)$\neg \forall x P(x)$

18. 假设命题函数 $Q(x)$的论域为整数-2，-1，0，1 和 2。使用析取、合取和否定写出下列命题。

a) $\exists x P(x)$　b) $\forall x P(x)$　c) $\exists x \neg P(x)$

d) $\forall x \neg P(x)$　e)$\neg \exists x P(x)$　f)$\neg \forall x P(x)$

19. 假设命题函数 $P(x)$的论域为整数 1，2，3，4 和 5。使用析取、合取和否定(而不使用量词)表达下列语句。

a) $\exists x P(x)$　b) $\forall x P(x)$　c)$\neg \exists x P(x)$　d)$\neg \forall x P(x)$

e) $\forall x((x \neq 3) \rightarrow P(x)) \vee \exists x \neg P(x)$

20. 假设命题函数 $P(x)$的论域为整数-5，-3，-1，1，3 和 5。使用析取、合取和否定(而不使用量词)表达下列语句。

a) $\exists x P(x)$　b) $\forall x P(x)$　c) $\forall x((x \neq 1) \rightarrow P(x))$

d) $\exists x((x \geqslant 0) \wedge P(x))$　e) $\exists x(\neg P(x)) \wedge \forall x((x<0) \rightarrow P(x))$

21. 使用谓词、量词和逻辑联接词，以两种方式将下列语句翻译成逻辑表达式。首先，令论域为班上的学

生，其次，令论域为所有人。

a)班上有人会说印地语。

b)班上的每个人都很友好。

c)班上有个学生不是出生在加利福尼亚。

d)班上有个学生曾演过电影。

e)班上没有学生上过逻辑编程课程。

22. 使用谓词、量词和逻辑联接词，以两种方式将下列语句翻译成逻辑表达式。首先，令论域为班上的学生，其次，令论域为所有人。

a)班上的每个学生都有移动电话。

b)班上的某个学生曾看过外国影片。

c)班上的某个学生不会游泳。

d)班上的所有学生都会求解二次方程。

e)班上的某个学生不想变富。

23. 使用谓词、量词和逻辑联接词，将下列语句翻译成逻辑表达式。

a)人无完人。

b)不是每个人都很完美。

c)你的所有朋友都很完美。

d)你的某个朋友很完美。

e)每个人都是你的朋友并且很完美。

f)不是每个人都是你的朋友或有人并不完美。

24. 使用谓词(带有一个或两个变量)，通过改变论域，以三种不同的方式将下列语句翻译成逻辑表达式。

a)学校中的某个人去过乌兹别克斯坦。

b)班上的每个人都学过微积分和 C++ 。

c)学校里没有人同时拥有摩托车和自行车。

d)学校里有某个人不快乐。

e)学校里的每个人都生于 20 世纪。

25. 使用谓词(带有一个或两个变量)，通过改变论域，以三种不同的方式将下列语句翻译成逻辑表达式。

a)学校里的某个学生曾在越南居住过。

b)学校里的有个学生不会说印地语。

c)学校里的某个学生会用 Java、Prolog 和 C++ 。

d)班上的每个学生都喜欢泰国食物。

e)班上的某个学生不打曲棍球。

26. 使用谓词、量词和逻辑联接词，将下列语句翻译成逻辑表达式。

a)某东西不在正确的位置。

b)所有的工具都在正确的位置并且状况良好。

c)每样东西都在正确的位置并且状况良好。

d)没有东西在正确的位置并且状况良好。

e)有某样东西不在正确的位置，但它状况良好。

27. 使用逻辑运算符、谓词和量词来表达下列语句。

a)某些命题是永真式。

b)矛盾的否定是永真式。

c)两个可能式的析取可以是一个永真式。

d)两个永真式的合取是一个永真式。

28. 假定命题函数 $P(x, y)$的论域由 x 和 y 的序偶组成，其中 x 是 1、2 或 3，y 是 1、2 或 3。用析取和合取写出下列命题。

a) $\exists x P(x, 3)$

b) $\forall y P(1, y)$

c) $\exists y \neg P(2, y)$

d) $\forall x \neg P(x, 2)$

29. 假定 $Q(x, y, z)$的论域由 x、y 和 z 的三元组组成，其中 x=0、1 或 2，y=0 或 1，z=0 或 1。用析取和合取写出下列命题。

a) $\forall y Q(0, y, 0)$

b) $\exists x Q(x, 1, 1)$

c) $\exists z \neg Q(0, 0, z)$

d) $\exists x \neg Q(x, 0, 1)$

30. 用量词表达下列语句，然后取该语句的否定并使否定词不在量词的左边。再用简单语句表达这否定(不要简单地表达为“不是……”)。

a)所有的狗都长跳蚤。

b)有一匹马会做加法。

c)每只考拉都会爬树。

d)没有会说法语的猴子。

e)有会游泳和捕鱼的猪。

31. 用量词表达下列语句，然后取该语句的否定并使否定词不在量词的左边。再用简单语句表达这否定(不要简单地表达为“不是……”)。

a)一些年长的狗会学习新的技巧。

b)没有会微积分的兔子。

c)每只鸟都会飞。

d)没有会说话的狗。

e)班上没有人会法语和俄语。

32. 用量词表达下列命题的否定，再用语句表达这些否定。

a)一些司机不遵守驾驶速度限制。

b)所有的瑞典电影都很严肃。

c)没人能保守秘密。

d)班上的某人的态度不好。

33. 找出下列全称量化语句的反例(如果可能的话)，其中所有变量的论域是整数集合。

a) $\forall x(x^2 \geqslant x)$　　b) $\forall x(x>0 \vee x<0)$　　c) $\forall x(x=1)$

34. 找出下列全称量化语句的反例(如果可能的话)，其中所有变量的论域是实数集合。

a) $\forall x(x^2 \neq x)$　　b) $\forall x(x^2 \neq 2)$　　c) $\forall x(|x|>0)$

35. 用谓词和量词表达下列语句。

a)一名飞机乘客如果在一年中飞行的里程超过25 000英里，或在那一年坐飞机的次数超过25次，就表明他是高贵乘客。

b)一名男选手，若他的马拉松比赛过去最好成绩在3小时内，他就有资格参赛。一名女选手，若她的过去最好成绩在3.5小时内，她就有资格参赛。

c)一名学生要想取得硕士学位，必须至少修满60个学分，或至少修满45个学分并通过硕士论文答辩，并且所有必修课程的成绩不低于B。

d)有某个学生在一个学期内修了21个学分并且全部课程的成绩都为A。

练习36～40主要进行系统规范说明和逻辑表达式之间的翻译，当中涉及量词。

36. 将下列系统规范说明翻译成语句，其中谓词$S(x, y)$是"x在状态y"，x和y的论域分别是所有系统和所有可能的状态。

a) $\exists x S(x, \text{open})$

b) $\forall x(S(x, \text{malfunctioning}) \vee S(x, \text{diagnostic}))$

c) $\exists x S(x, \text{open}) \vee \exists x S(x, \text{diagnostic})$

d) $\exists x \neg S(x, \text{available})$

e) $\forall x \neg S(x, \text{working})$

37. 将下列规范说明翻译成语句，其中$F(p)$是"打印机p不能提供服务"，$B(p)$是"打印机p很忙"，$L(j)$是"打印作业j丢失了"，$Q(j)$是"打印作业j在队列中"。

a) $\exists p(F(p) \wedge B(p)) \rightarrow \exists j L(j)$

b) $\forall p B(p) \rightarrow \exists j Q(j)$

c) $\exists j(Q(j) \wedge L(j)) \rightarrow \exists p F(p)$

d) $(\forall p B(p) \wedge \forall j Q(j)) \rightarrow \exists j L(j)$

38. 使用谓词、量词和逻辑联接词表达下列系统规范说明。

a)当硬盘中的空闲空间少于30MB时，就会向所有用户发送警告消息。

b)当检测到系统错误时，文件系统中的目录均不能打开且不能关闭文件。

c)如果当前有用户登录了，就不能备份文件系统。

d)当有至少8MB内存可用且连接速度至少为56Kbps时，就可以进行视频点播。

39. 使用谓词、量词和逻辑联接词表达下列系统规范说明。

a)如果磁盘有10MB以上的空闲空间，那么至少能保存一条邮件消息。

b)每当有主动报警时，所有排队的消息都被传送。

c)诊断监控器跟踪所有系统的状态，除了主控制台。

d)对参与电话会议的每一方，不在特殊列表上的主叫方应当付费。

40. 使用谓词、量词和逻辑联接词表达下列系统规范说明。

a)每个用户都可以访问电子邮箱。

b)如果文件系统被锁定，组中的每个人都能访问系统邮箱。

c)防火墙处于诊断状态仅当代理服务器处于诊断状态。

d)如果吞吐量在100kbps和500kbps之间且代理服务器不处于诊断模式，则至少有一个路由器工作

正常。

41. 判断 $\forall x(P(x)\rightarrow Q(x))$ 和 $\forall xP(x)\rightarrow\forall xQ(x)$ 是否具有相同的真值。
42. 证明语句 $\forall x(P(x)\wedge Q(x))$ 和 $\forall xP(x)\wedge\forall xQ(x)$ 有同样的真值。
43. 证明 $\exists x(P(x)\vee Q(x))$ 和 $\exists xP(x)\vee\exists xQ(x)$ 有同样的真值。
44. 证明下列逻辑等价关系，其中 A 是不含量词的命题。

    a) $(\forall xP(x))\vee A\Leftrightarrow\forall x(P(x)\vee A)$　　b) $(\exists xP(x))\vee A\Leftrightarrow\exists x(P(x)\vee A)$
45. 证明下列逻辑等价关系，其中 A 是不含量词的命题。

    a) $(\forall xP(x))\wedge A\Leftrightarrow\forall x(P(x)\wedge A)$　　b) $(\exists xP(x))\wedge A\Leftrightarrow\exists x(P(x)\wedge A)$
46. 证明 $\forall xP(x)\vee\forall xQ(x)$ 和 $\forall x(P(x)\vee Q(x))$ 不是逻辑等价的。
47. 证明 $\exists xP(x)\wedge\exists xQ(x)$ 和 $\exists x(P(x)\wedge Q(x))$ 不是逻辑等价的。
48. 符号 $\exists!xP(x)$ 表示命题

    “有唯一的 x 使 $P(x)$ 成真。”

    论域是整数集合，下列各项的真值是什么？

    a) $\exists!x(x>1)$　　b) $\exists!x(x^2=1)$　　c) $\exists!x(x+3=2x)$　　d) $\exists!x(x=x+1)$
49. 下列语句的真值是什么？

    a) $\exists!xP(x)\rightarrow\exists xP(x)$　b) $\forall xP(x)\rightarrow\exists!xP(x)$　c) $\exists!x\neg P(x)\rightarrow\neg\forall xP(x)$
50. 假定论域是 1、2 和 3，用否定、合取和析取写出量化语句 $\exists!xP(x)$。
51. 给定例 21 的 Prolog 事实，对下列查询 Prolog 返回的是什么？

    a)`? instructor(chan, math273)`　　b)`? instructor(patel, cs301)`

    c)`? enrolled(X, cs301)`　　d)`? enrolled(kiko, Y)`

    e)`? teaches(grossman, Y)`
52. 给定例 21 的 Prolog 事实，对下列查询 Prolog 返回的是什么？

    a)`? enrolled(kevin, ee222)`　　b)`? enrolled(kiko, math273)`

    c)`? instructor(grossman, X)`　　d)`? instructor(X, cs301)`

    e)`? teaches(X, kevin)`
53. 假定 Prolog 事实用于定义谓词 *mother*(M, Y) 和 *father*(F, X)，这表示 M 是 Y 的母亲，F 是 X 的父亲。给出一个 Prolog 规则定义谓词 *sibling*(X, Y)，表示 X 和 Y 是兄弟(也就是，有相同的父亲和母亲)。
54. 假定 Prolog 事实用于定义谓词 *mother*(M, Y) 和 *father*(F, X)，这表示 M 是 Y 的母亲，F 是 X 的父亲。给出一个 Prolog 规则定义谓词 *grandfather*(X, Y)，表示 X 是 Y 的祖父。(提示：可以在 Prolog 中写一个析取，使用分号分开谓词或将谓词放在单独的行中。)

练习 55～58 是根据 Lewis Carroll 的《*Symbolic Logic*》一书中的问题编写的。

55. 令 $P(x)$、$Q(x)$ 和 $R(x)$ 分别为语句“x 是教授”、“x 无知”和“x 爱虚荣”。用量词、逻辑联接词和 $P(x)$、$Q(x)$、$R(x)$ 表达下列语句。假定论域是所有人的集合。

    a)没有无知的教授。

    b)所有无知者均爱虚荣。

    c)没有爱虚荣的教授。

    d)能从 a)和 b)推出 c)吗？若不能，有没有一个正确的结论？
56. 令 $P(x)$、$Q(x)$ 和 $R(x)$ 分别表示语句“x 是个清楚的解释”、“x 令人满意”和“x 是借口”。假定 x 的论域是所有英语文章。用量词、逻辑联接词和 $P(x)$、$Q(x)$、$R(x)$ 表达下列语句。

    a)所有清楚的解释都令人满意。

b)有些借口令人不满意。

c)有些借口不是清楚的解释。

*d)能从a)和b)推出c)吗？若不能，有没有一个正确的结论？

57. 令 $P(x)$、$Q(x)$、$R(x)$ 和 $S(x)$ 分别为语句“x 是婴儿”、“x 的行为符合逻辑”、“x 能管理鳄鱼”和“x 被人轻视”。假定 x 的论域是所有人的集合。用量词、逻辑联接词和 $P(x)$、$Q(x)$、$R(x)$ 和 $S(x)$ 表达下列语句。

a)婴儿行为不合逻辑。

b)能管理鳄鱼的人不被人轻视。

c)行为不合逻辑的人被人轻视。

d)婴儿不能管理鳄鱼。

*e)能从a)、b)和c)推出d)吗？若不能，有没有一个正确的结论？

58. 令 $P(x)$、$Q(x)$、$R(x)$ 和 $S(x)$ 分别为语句“x 是只鸭子”、“x 是我的一只家禽”、“x 是一名军官”和“x 愿意跳华尔兹”。用量词、逻辑联接词和 $P(x)$、$Q(x)$、$R(x)$ 和 $S(x)$ 表达下列语句。

a)没有鸭子愿意跳华尔兹。

b)没有军官不会跳华尔兹。

c)我所有的家禽都是鸭子。

d)我的家禽都不是军官。

*e)能从a)、b)和c)推出d)吗？若不能，有没有一个正确的结论？

1.4 嵌套量词

1.4.1 引言

在1.3节我们定义了存在量词和全称量词，并展示了如何用它们来表示数学语句。我们也解释了如何用它们来将汉语语句翻译成逻辑表达式。本节我们将学习嵌套量词，它是出现在其他量词的作用域内的量词，如在语句 $\forall x \exists y(x+y=0)$ 中。嵌套量词通常在数学和计算机科学中出现。尽管嵌套量词有时比较难理解，但在1.3节介绍过的规则却有助于我们使用它们。

1.4.2 翻译涉及嵌套量词的语句

复杂的表达式往往会涉及在各种上下文中出现的量词。要理解这些涉及许多量词的语句，我们需要阐明出现的量词和谓词的含义。具体如例1所示。

例1 假定变量 x 和 y 的论域是所有实数的集合，语句

Additional Steps

$$\forall x \forall y(x+y=y+x)$$

说的是对所有实数 x 和 y，$x+y=y+x$。这是实数加法的交换律。同样，语句

$$\forall x \exists y(x+y=0)$$

Extra Examples

说的是对所有实数 x，有一个实数 y，使得 $x+y=0$。也就是每个实数都有一个加法的逆。同样，语句

$$\forall x \forall y \forall z(x+(y+z)=(x+y)+z)$$

是实数加法的结合律 ■

例2 将下列语句翻译成汉语句子

$$\forall x \forall y((x>0) \wedge (y<0) \rightarrow (xy<0))$$

其中变量 x 和 y 的论域都是全体实数。

解 这个语句表示对任意实数 x 和 y，如果 $x>0$ 且 $y<0$，那么 $xy<0$。也就是说，这个语句表示对实数 x 和 y，如果 x 是正的且 y 是负的，那么 xy 就是负的。这可以更简洁地叙述为“一个正实数与一个负实数的积是一个负实数”。 ■

表达汉语语句的带嵌套量词的表达式可以相当复杂。在翻译这样的表达式时，第一步是写

出表达式中的量词和谓词的含义，第二步是用简单的句子来表达这个含义。例 3 和例 4 示出了这个过程。

例 3 把语句

$$\forall x(C(x) \vee \exists y(C(y) \wedge F(x,y)))$$

译成文字语句，其中 $C(x)$是“x 有台计算机”，$F(x, y)$是“x 和 y 是朋友”，而 x 和 y 的共同论域是学校全体学生的集合。

解 该语句说的是，对学校的每一个学生 x，或者 x 有台计算机，或者另有学生 y，他有台计算机，并且 x 和 y 是朋友。换言之，学校的每个学生或有计算机，或有个有计算机的朋友。■

例 4 把语句

$$\exists x \forall y \forall z((F(x,y) \wedge F(x,z) \wedge (y \neq z)) \rightarrow \neg F(y,z))$$

译成文字语句，其中 $F(a, b)$的含义是 a 和 b 是朋友，而 x、y 和 z 的论域是学校所有学生的集合。

解 这个语句说的是，有一个学生 x，对所有学生 y 及不同于 y 的所有学生 z，只要 x 和 y 是朋友，x 和 z 也是朋友，那么 y 和 z 就不是朋友。换句话说，有个学生，他的朋友之间都不是朋友。■

1.4.3 将语句翻译成逻辑表达式

在 1.3 节我们展示了如何用量词来将句子翻译成逻辑表达式。然而，当时没有谈及这样的句子，即将其翻译成逻辑表达式要求使用嵌套量词。我们现在介绍如何翻译这样的句子。

例 5 将语句“如果某人是女性且为人父母，这个人是某人的母亲”翻译成逻辑表达式，当中涉及谓词、量词(论域是所有人)以及逻辑联接词。

解 语句“如果某人是女性且为人父母，这个人是某人的母亲”可以表达为“对每个人 x，如果 x 是女性且 x 是父母，那么存在一个人 y 且 x 是 y 的母亲”。我们引入谓词 $F(x)$来表示“x 是女性”，$P(x)$表示“x 是父母”，$M(x, y)$表示“x 是 y 的母亲”。则初始语句可以表示为

$$\forall x((F(x) \wedge P(x)) \rightarrow \exists y M(x,y))$$

我们可以把 $\exists y$ 移到左边，因为 y 不在 $F(x) \wedge P(x)$中出现，这样得到一个等价的表达式

$$\forall x \exists y((F(x) \wedge P(x)) \rightarrow M(x,y))$$

■

例 6 将语句“每个人都恰有一个最好的朋友”翻译成逻辑表达式，当中涉及谓词、量词(论域是所有人)以及逻辑联接词。

解 语句“每个人都恰有一个最好的朋友”可以表达为“对每个人 x，x 恰有一个最好的朋友”。引入全称量词，可以看到这个语句等价于“$\forall x$(x 恰有一个最好的朋友)”，其中论域是所有人。

x 恰有一个最好的朋友意味着有一个人 y 是 x 的最好的朋友，而且，对每个人 z，如果 z 不是 y，那么 z 不是 x 的最好的朋友。当我们引入谓词 $B(x, y)$为语句“y 是 x 的最好的朋友”，则语句“x 恰有一个最好的朋友”可以表示为

$$\exists y(B(x,y) \wedge \forall z((z \neq y) \rightarrow \neg B(x,z)))$$

因此，初始语句可以表示为

$$\forall x \exists y(B(x,y) \wedge \forall z((z \neq y) \rightarrow \neg B(x,z)))$$

■

例 7 用量词表示语句“有位妇女已搭乘过世界上每一条航线上的航班”。

解 令 $P(w, f)$为“w 搭乘过 f”，而 $Q(f, a)$为“f 是 a 上的航班”。于是可将上述语句表示为

$$\exists w \forall a \exists f(P(w,f) \wedge Q(f,a))$$

其中，w、f 和 a 的论域分别为世界上所有妇女、所有空中航班和所有航线。

这个语句也可以表示为

$$\exists w \forall a \exists f R(w,f,a)$$

其中 $R(w, f, a)$为“w 已搭乘过 a 上的 f”。虽然这样表示更紧凑，但它使变量之间的关系有点含糊不清，因此，第一个解要好些。■

用汉语表达的数学语句可以被翻译成逻辑表达式，如例 8 至例 10 所示。

例 8 将语句“两个正整数的和是正数”翻译成逻辑表达式。

Additional Steps

解 要将这个语句翻译成逻辑表达式，我们首先重写它以示出隐含的量词“对每两个正整数，它们的和是正数”。然后，引入变量 x 和 y 就得到“对所有正整数 x 和 y，$x+y$ 是正数”。因此，可以将这个语句表达为

$$\forall x \forall y((x > 0) \wedge (y > 0) \rightarrow (x + y > 0))$$

其中这两个变量的论域是全体整数。■

例 9 将语句“每个实数(除了零)都有*乘法逆元*”翻译成逻辑表达式。

Extra Examples

解 我们首先重写这个语句为“对每个实数 x(除了零)，x 有乘法逆元”，然后可以再将之重写为“对每个实数 x，如果 $x\neq 0$，那么存在一个实数 y 使得 $xy=1$”。这可以重写为

$$\forall x((x \neq 0) \rightarrow \exists y(xy = 1))$$ ■

正如前面提到的，量词常用于数学概念的定义，你可能熟悉的极限概念即为一例。极限是微积分中的重要概念。

例 10 **（需要微积分知识）** 用量词表示极限的定义。

解 回顾一下，语句$\lim_{x\to a} f(x)=L$ 的定义是：对每个实数 $\epsilon>0$，存在一个实数 $\delta>0$，使得对每一个 x，只要 $0<|x-a|<\delta$，就有 $|f(x)-L|<\epsilon$。极限的这一定义用量词表示为

$$\forall \epsilon \exists \delta \forall x(0 < |x - a| < \delta \rightarrow |f(x) - L| < \epsilon)$$

其中 ϵ 和 δ 的论域是正实数集合，x 的论域是实数集合。

这一定义还可表示为

$$\forall \epsilon > 0 \exists \delta > 0 \forall x(0 < |x - a| < \delta \rightarrow |f(x) - L| < \epsilon)$$

其中 ϵ 和 δ 的论域为实数集合，而不是正实数集合。■

1.4.4 否定嵌套量词

带嵌套量词的语句可以通过连续地应用否定带单个量词的语句的规则成为否定的。具体如例 11～例 13 所示。

例 11 表达语句 $\forall x \exists y(xy=1)$的否定，使得量词前面没有否定词。

Extra Examples

解 通过连续地应用否定量化语句的规则(见 1.3 节表 1－17)，可以将 $\neg\forall x \exists y(xy=1)$ 中的否定词移入所有量词里面。我们发现，$\neg\forall x \exists y(xy=1)$等价于 $\exists x \neg\exists y(xy=1)$，而后者又等价于 $\exists x \forall y \neg(xy=1)$。由于 $\neg(xy=1)$可以简化为 $xy\neq 1$，所以，我们的否定语句可以表达为 $\exists x \forall y(xy\neq 1)$。■

例 12 使用量词表达语句“没有一个妇女已搭乘过世界上每一条航线上的航班”。

解 这个语句是例 7 的语句“有位妇女已搭乘过世界上每一条航线上的航班”的否定。由例 7，我们的语句可以表达为 $\neg\exists w \forall a \exists f(P(w, f) \wedge Q(f, a))$，其中 $P(w, f)$是“w 搭乘过 f”，

而 $Q(f, a)$为"f 是 a 上的航班"。通过连续地应用否定量化语句的规则(见 1.3 节表 1-17)并在最后一步应用德摩根律，可发现我们的语句等价于如下语句：

$$\begin{aligned}\forall w \neg \forall a \exists f(P(w,f) \wedge Q(f,a)) &\equiv \forall w \exists a \neg \exists f(P(w,f) \wedge Q(f,a)) \\ &\equiv \forall w \exists a \forall f \neg (P(w,f) \wedge Q(f,a)) \\ &\equiv \forall w \exists a \forall f(\neg P(w,f) \vee \neg Q(f,a))\end{aligned}$$ ■

例 13 使用量词和谓词表达$\lim_{x \to a} f(x)$不存在这一事实。

解 $\lim_{x \to a} f(x)$不存在意味着对全体实数 L，$\lim_{x \to a} f(x) \neq L$。根据例 10，$\lim_{x \to a} f(x) \neq L$ 可以表达为

$$\neg \forall \varepsilon > 0 \exists \delta > 0 \forall x(0 < |x-a| < \delta \rightarrow |f(x)-L| < \varepsilon)$$

连续地应用否定量化表达式的规则，可得到如下等价语句序列：

$$\begin{aligned}&\neg \forall \varepsilon > 0 \exists \delta > 0 \forall x(0 < |x-a| < \delta \rightarrow |f(x)-L| < \varepsilon) \\ &\equiv \exists \varepsilon > 0 \neg \exists \delta > 0 \forall x(0 < |x-a| < \delta \rightarrow |f(x)-L| < \varepsilon) \\ &\equiv \exists \varepsilon > 0 \forall \delta > 0 \neg \forall x(0 < |x-a| < \delta \rightarrow |f(x)-L| < \varepsilon) \\ &\equiv \exists \varepsilon > 0 \forall \delta > 0 \exists x \neg (0 < |x-a| < \delta \rightarrow |f(x)-L| < \varepsilon) \\ &\equiv \exists \varepsilon > 0 \forall \delta > 0 \exists x(0 < |x-a| < \delta \wedge |f(x)-L| \geqslant \varepsilon)\end{aligned}$$

在最后一步使用等价式 $\neg (p \rightarrow q) \equiv p \wedge \neg q$ 。

由于$\lim_{x \to a} f(x)$不存在意味着对全体实数 L，$\lim_{x \to a} f(x) \neq L$。这个语句可以表达为

$$\forall L \exists \varepsilon > 0 \forall \delta > \exists x(0 < |x-a| < \delta \wedge |f(x)-L| \geqslant \varepsilon).$$

这最后一个语句表示，对每个实数 L，存在实数 $\varepsilon>0$ 使得对每个实数 $\delta>0$，存在实数 x 使得 $0<|x-a|<\delta$和 $|f(x)-L| \geqslant \varepsilon$。 ■

1.4.5 量词的顺序

Extra Examples 许多数学语句需要对多变量命题函数作多重量化。除非所有量词均为全称量词或均为存在量词，否则量词的顺序是重要的。例 14～16 用于说明这一点。这几个例子中每个变量的论域都假定为实数集合。

例 14 令 $P(x, y)$为语句"$x+y=y+x$"，量化语句 $\forall x \forall y P(x, y)$的真值是什么？

解 量化语句

$$\forall x \forall y P(x,y)$$

表示的命题是

"对所有实数 x 和所有实数 y, $x+y=y+x$ 成立。"

因为 $P(x, y)$对所有实数 x 和 y 成真，$\forall x \forall y P(x, y)$成真。 ■

例 15 令 $Q(x, y)$表示"$x+y=0$"，量化语句 $\exists y \forall x Q(x, y)$和 $\forall x \exists y Q(x, y)$的真值是什么？

解 量化语句

$$\exists y \forall x Q(x,y)$$

表示的命题是

"有个实数 y 能使 $Q(x,y)$ 对每一个实数 x 成立。"

不管 y 取什么值，只有一个 x 的值能使 $x+y=0$ 成立。因为没有实数 y 能使 $x+y=0$ 对所有实数 x 成立，语句 $\exists y \forall x Q(x, y)$为假。

量化语句

$$\forall x \exists y Q(x,y)$$

表示的命题是

"对每个实数 x 都有一个实数 y 使 $Q(x, y)$ 成立。"
给定一个实数 x，总有一个实数 y 能使 $x+y=0$，这个实数就是 $y=-x$。因此，语句 $\forall x \exists y Q(x, y)$ 为真。 ■

例 15 说明量词出现的顺序不同含义也不同。语句 $\exists y \forall x P(x, y)$ 和 $\forall x \exists y P(x, y)$ 不是逻辑等价的，语句 $\exists y \forall x P(x, y)$ 成真当且仅当存在一个 y，使得 $P(x, y)$ 对每个 x 都成立。因此，要使这一语句为真，必须有一个特定的 y 值，对这个值无论是什么样的 x，$P(x, y)$ 都成立；另一方面，$\forall x \exists y P(x, y)$ 为真当且仅当对 x 的每一个值都有一个 y 的值使 $P(x, y)$ 成立。所以，要使这个语句为真，不管你选什么 x，总有 y 的一个值(也许依赖于你选的 x)使 $P(x, y)$ 成立。换言之，在第二种情况下，y 随着 x 而定，但在第一种情况下，y 是与 x 无关的常数。

从这些观察可以看出，如果 $\exists y \forall x P(x, y)$ 为真，则 $\forall x \exists y P(x, y)$ 必定也为真；可是如果 $\forall x \exists y P(x, y)$ 为真，$\exists y \forall x P(x, y)$ 不一定为真。(参见本章补充练习 14 和 16。)

循环量化思考 对多个变量使用量词时，借助嵌套循环来思考是有益的。(当然，如果某个变量的论域有无穷多个元素，是无法真正对所有值做循环的。不过这种思考方式对理解嵌套量词总是有益的。)例如，要决定 $\forall x \forall y P(x, y)$ 是否为真，我们先对 x 的所有值做循环，而对 x 的每个值再对 y 的所有值循环。如果我们发现对 x 和 y 的所有值 $P(x, y)$ 都为真，那么我们就判定了 $\forall x \forall y P(x, y)$ 为真。只要我们碰上一个 x 值，对这个值又有一个 y 值使 $P(x, y)$ 为假，那么就证明了 $\forall x \forall y P(x, y)$ 为假。

同样，要判定 $\forall x \exists y P(x, y)$ 是否为真，我们对 x 的所有值循环，对 x 的每个值，对 y 的值循环直到找到一个 y 使 $P(x, y)$ 为真。如果对 x 的所有值，我们都能碰上这样的一个 y 值，那么 $\forall x \exists y P(x, y)$ 为真。如果对某个 x 我们碰不上这样的 y，那么 $\forall x \exists y P(x, y)$ 就为假。

要决定 $\exists x \forall y P(x, y)$ 是否为真，需要对 x 的值循环直到找到某个 x，就这个 x 对 y 的所有值循环时 $P(x, y)$ 总是成真。如果能找到这样的 x，$\exists x \forall y P(x, y)$ 就为真。如果总也碰不上这样的 x，那么我们知道 $\exists x \forall y P(x, y)$ 为假。

最后要看 $\exists x \exists y P(x, y)$ 是否为真。我们对 x 的值循环，循环时对 x 的每个值都对 y 的值循环，直到找到 x 的一个值和 y 的一个值使 $P(x, y)$ 为真。只有当我们永远碰不上这样的 x 和 y 能使 $P(x, y)$ 成真时，语句 $\exists x \exists y P(x, y)$ 才为假。

表 1-18 是两个变量所有不同的可能量化方式的总结。

例 16 说明，两个以上变量的量化也是常见的。

表 1-18 两个变量的量化

语　　句	何时为真?	何时为假?
$\forall x \forall y P(x, y)$ $\forall y \forall x P(x, y)$	对每一对 x、y，$P(x, y)$ 均为真	有一对 x、y 使 $P(x, y)$ 为假
$\forall x \exists y P(x, y)$	对每个 x，都有 y 使 $P(x, y)$ 为真	有 x，使 $P(x, y)$ 对每个 y 总是假
$\exists x \forall y P(x, y)$	有一个 x，使 $P(x, y)$ 对所有 y 均为真	对每个 x 都有 y 使 $P(x, y)$ 为假
$\exists x \exists y P(x, y)$ $\exists y \exists x P(x, y)$	有一对 x、y 使 $P(x, y)$ 为真	对每一对 x、y，$P(x, y)$ 均为假

例 16 令 $Q(x, y, z)$ 为语句"$x+y=z$"，语句 $\forall x \forall y \exists z Q(x, y, z)$ 和 $\exists z \forall x \forall y Q(x, y, z)$ 的真值是什么?

解 假定给 x 和 y 赋了值，那么就有一个实数 z，使得 $x+y=z$。于是量化语句

$$\forall x \forall y \exists z Q(x,y,z)$$

为真，因为它其实就是语句

"对所有实数 x 和所有实数 y，有实数 z，使得 $x+y=z$"。

这里量词出现的顺序是重要的，因为量化语句

$$\exists z \forall x \forall y Q(x,y,z)$$

也就是语句

"有实数 z 使得对所有实数 x 和所有实数 y，$x+y=z$"

为假，因为没有 z 的值能使 $x+y=z$ 对 x 和 y 的所有值都成立。 ■

练习

1. 将下列语句翻译成汉语句子，其中每个变量的论域是全体实数。
   a) $\forall x \exists y(x<y)$　　b) $\forall x \forall y(((x\geqslant 0)\wedge(y\geqslant 0))\rightarrow(xy\geqslant 0))$
   c) $\forall x \forall y \exists z(xy=z)$
2. 将下列语句翻译成汉语句子，其中每个变量的论域是全体实数。
   a) $\exists x \forall y(xy=y)$　　b) $\forall x \forall y(((x\geqslant 0)\wedge(y<0))\rightarrow(x-y>0))$
   c) $\forall x \forall y \exists z(x=y+z)$
3. 令 $Q(x, y)$是语句"x 已经发送电子邮件消息给 y"，其中 x 和 y 的论域都是班上的所有学生，将下列量化式子表达成汉语句子。
   a) $\exists x \exists y Q(x, y)$　　b) $\exists x \forall y Q(x, y)$　　c) $\forall x \exists y Q(x, y)$
   d) $\exists y \forall x Q(x, y)$　　e) $\forall y \exists x Q(x, y)$　　f) $\forall x \forall y Q(x, y)$
4. 令 $P(x, y)$表示语句"x 选修 y"，其中 x 的论域是班上全体学生的集合，y 的论域是你校所有计算机科学课程的集合。用句子表达下列各量化语句。
   a) $\exists x \exists y P(x, y)$　　b) $\exists x \forall y P(x, y)$　　c) $\forall x \exists y P(x, y)$
   d) $\exists y \forall x P(x, y)$　　e) $\forall y \exists x P(x, y)$　　f) $\forall x \forall y P(x, y)$
5. 令 $W(x, y)$表示"x 访问过 y"，其中 x 的论域是你校全体学生集合，y 的论域是所有网站的集合。用简单的句子表达下列语句。
   a) W(Sarah Smith, www. att. com)
   b) $\exists x W(x$, www. imdb. org)
   c) $\exists y W$(Jose Orez, y)
   d) $\exists y(W$(Ashok Puri, y) $\wedge$ W(Cindy Yoon, y))
   e) $\exists y \forall z(y\neq$(David Belcher) $\wedge$ (W(David Belcher, z) $\rightarrow W(y, z)$))
   f) $\exists x \exists y \forall z((x\neq y)\wedge(W(x, z)\leftrightarrow W(y, z)))$
6. 令 $C(x, y)$表示"x 注册了 y"，其中 x 的论域是你校全体学生的集合，y 的论域是你校开设所有课程的集合。用简单的句子表达下列语句。
   a) C(Randy Goldberg, CS 252)　　b) $\exists x C(x$, Math 695)
   c) $\exists y C$(Carol Sitea, y)　　d) $\exists x(C(x$, Math 222) $\wedge$ $C(x$, CS 252))
   e) $\exists x \exists y \forall z((x\neq y)\wedge(C(x, z)\rightarrow C(y, z)))$　　f) $\exists x \exists y \forall z((x\neq y)\wedge(C(x, z)\leftrightarrow C(y, z)))$
7. 令 $T(x, y)$表示学生 x 喜欢菜肴 y，其中 x 的论域是学校的所有学生，y 的论域是所有菜肴。用简单的汉语句子表达下列语句。
   a) $\neg$ T(Abdallah Hussein, Japanese)　　b) $\exists x T(x$, Korean) $\wedge$ $\forall x T(x$, Mexican)
   c) $\exists y(T$(Monique Arsenault, y) $\vee$ T(Jay Johnson, y))　　d) $\forall x \forall z \exists y((x\neq z)\rightarrow\neg(T(x, y)\wedge T(z, y)))$
   e) $\exists x \exists z \forall y(T(x, y)\leftrightarrow T(z, y))$　　f) $\forall x \forall z \exists y(T(x, y)\leftrightarrow T(z, y))$
8. 令 $Q(x, y)$为语句"x 为 y 的参赛者"。用 $Q(x, y)$、量词和逻辑联接词表达下列各句，其中 x 的论域是你校所有学生的集合，y 的论域是所有电视智力竞赛节目。
   a)你校有位学生参加了一个电视智力竞赛节目。
   b)你校没有学生参加过电视智力竞赛。

c)你校有位学生参加了 Jeopardy 和 Wheel of Fortune 两个电视竞赛。

d)每个电视智力竞赛节目都有你校的一名参赛学生。

e)你校至少有两个学生参加了 Jeopardy 比赛。

9. 令 $L(x, y)$为语句“x 爱 y”，其中 x 和 y 的论域都是全世界所有人的集合。用量词表达下列各语句。

a)每个人都爱 Jerry。
b)每个人都爱某个人。
c)有个人人都爱的人。
d)没有人爱所有的人。
e)有个 Lydia 不爱的人。
f)有个人人都不爱的人。
g)恰有一个人人都爱的人。
h)Lynn 爱的人恰有两个。
i)每个人都爱自己。
j)有人除自己以外谁都不爱。

10. 令 $F(x, y)$为语句“x 能骗 y”，其中 x 和 y 的论域为全世界所有人的集合。用量词表达下列语句。

a)人人都能骗 Fred。
b)Evelyn 能骗所有人。
c)每个人都能骗某个人。
d)没有人能骗所有人。
e)人人都会被人骗。
f)没有人能骗 Fred 和 Jerry 两个人。
g)Nancy 恰能骗两个人。
h)恰有一个人人都能骗的人。
i)没有人能骗自己。
j)有人除自己以外只能骗一个人。

11. 令 $S(x)$为谓词“x 是学生”，$F(x)$为谓词“x 是教员”，而 $A(x, y)$是谓词“x 问过 y 问题”，其中论域是你校所有人员的集合。用量词表达下列语句。

a)Lois 问过 Michaels 教授问题。
b)每个学生都问过 Gross 教授问题。
c)每位教员都问过 Miller 教授问题或被 Miller 教授问过问题。
d)某个学生从未问过任何教员问题。
e)有位教员从未被学生问过问题。
f)有个学生问过所有教员问题。
g)有位教员问过所有其他教员问题。
h)有学生从未被教员问过问题。

12. 令 $I(x)$为语句“x 能上因特网”，$C(x, y)$为语句“x 和 y 在因特网上交谈过”，其中 x 和 y 的论域是你们班上所有学生的集合。用量词表达下列语句。

a)Jerry 没有上过因特网。
b)Rachel 没在因特网上与 Chelsea 交谈过。
c)Jan 和 Sharon 从未在因特网上交谈过。
d)班上没有人与 Bob 交谈过。
e)除 Joseph 以外，Sanjay 与每个人都交谈过。
f)班上某人没有上过因特网。
g)班上并非人人都上过因特网。
h)班上恰有一人上过因特网。
i)班上除一个学生外都上过因特网。
j)班上上因特网的人在因特网上与班上至少另一名学生交谈过。
k)班上有人上过因特网，但从未与班上其他人交谈过。
l)班上有两个学生没做过网上交谈。
m)班上有个学生与班上每个人都做过网上交谈。
n)班上至少有两个学生没有与同一个人做过网上交谈。
o)班上有两个学生，他们当中有一个与班上其余每个人都交谈过。

13. 令 $M(x, y)$为“x 已发给 y 电子邮件”，$T(x, y)$为“x 给 y 打过电话”，其中论域为你们班上所有学生。用量词表达下列语句。(假定所有被发出的电子邮件都能收到，尽管事实并非如此。)

a)Chou 从未给 Koko 发过电子邮件。
b)Arlene 从未给 Sarah 发过电子邮件，或未打过电话。
c)Jose 从未收到过 Deborah 的电子邮件。
d)班上每个学生都给 Ken 发过电子邮件。
e)班上没有人给 Nina 打过电话。
f)班上每个人或给 Avi 打过电话或给他发过电子邮件。
g)班上有个学生给班上其他人都发过电子邮件。

h)班上有人给班上其他人或打过电话，或发过电子邮件。

i)班上有两个学生互发过电子邮件。

j)班上有个学生给自己发过电子邮件。

k)班上有个学生既没从班上其他人那里收到过电子邮件，也没有人给他打过电话。

l)班上每个学生都或从同班同学那里收到过电子邮件，或接到过同班同学的电话。

m)班上至少有两个学生，一个给另一个发过电子邮件，第二个则给第一个打过电话。

n)班上有两个同学，他们当中有一个给班上其余同学或发过电子邮件，或打过电话。

14. 用量词和谓词表达下列语句。

a)班上有个学生会说印地语。

b)班上每个学生都会开车。

c)班上某学生去过阿拉斯加，但还没去过夏威夷。

d)班上所有学生都至少学过一种程序语言。

e)班上有个学生已选修学校某个系开设的所有课程。

f)班上有个学生恰好与一个同班同学在同一座城市长大。

g)班上每个学生都至少与一位同学至少在一个在线聊天室交谈过。

15. 用量词和谓词表达下列语句。

a)每位学计算机科学的学生都需要上一门离散数学课。

b)班上有个学生有一台个人电脑。

c)班上每个学生至少选修了一门计算机科学课。

d)班上有个学生至少选修过一门计算机科学课。

e)班上每个学生都去过校园里每座建筑。

f)班上有个学生去过校园里至少一座楼的每个房间。

g)班上每个学生都去过校园里每座楼的至少一个房间。

16. 离散数学班上有1个主修数学的新生，12个主修数学的二年级学生，15个主修计算机科学的二年级学生，2个主修数学的三年级学生和1个主修计算机科学的四年级学生。用量词表达下列语句，再给出其真值。

a)班上有个三年级学生。

b)班上每个学生都主修计算机科学。

c)班上有个学生既不主修数学，也不是三年级学生。

d)班上每个学生要么是二年级学生，要么主修计算机科学。

e)有一门主修课，有一名学生在每一学年都主修它。

17. 使用谓词、量词和逻辑联接词(如果有必要)表达下列系统规范说明。

a)每个用户恰能访问一个邮箱。

b)在所有错误条件下某个进程能继续运行，仅当内核运行正确。

c)校园网的所有用户都能访问具有.edu后缀的URL的所有站点。

*d)恰有两个系统监控着每一个远程服务器。

18. 使用谓词、量词和逻辑联接词(如果有必要)表达下列系统规范说明。

a)在各种故障情形下，至少要有一个控制台可以访问。

b)只要档案文件包含有至少一条由系统的用户发送的消息，就可以检索到用户的e-mail地址。

c)对每个安全漏洞，至少有一个机制可以检测到这个漏洞当且仅当有一个进程还未被损害。

d)对网络上的任意两个端点，至少有两条路径可以连接它们。

e)没有人知道系统的用户的口令，除了系统管理员。

19. 使用数学和逻辑运算符、谓词及量词表达下列语句，论域是全体整数。

a)两个负整数的和是负数。

b)两个正整数的差不一定是正数。

c)两个整数的平方和大于等于它们的和的平方。

d)两个整数的积的绝对值等于它们的绝对值的积。

20. 使用数学和逻辑运算符、谓词及量词表达下列语句，论域是全体整数。

a)两个负整数的积是正数。

b)两个正整数的平均数是正数。

c)两个负整数的差不一定是负数。

d)两个整数的和的绝对值不超过它们的绝对值的和。

21. 使用谓词、量词、逻辑联接词和数学运算符表达语句"每个正整数是四个整数的平方的和"。

22. 使用谓词、量词、逻辑联接词和数学运算符表达语句"有某个正整数不是三个整数的平方的和"。

23. 使用谓词、量词、逻辑联接词和数学运算符表达下列数学语句：

a)两个负实数的积是正数。

b)一个实数与它自身的差是零。

c)每个正实数恰有两个平方根。

d)负实数没有实数平方根。

24. 将下列嵌套量化式子翻译成表达出数学事实的汉语语句。论域均为全体实数。

a) $\exists x \forall y(x+y=y)$

b) $\forall x \forall y(((x\geqslant 0)\wedge(y<0))\rightarrow(x-y>0))$

c) $\exists x \exists y(((x\leqslant 0)\wedge(y\leqslant 0))\wedge(x-y>0))$

d) $\forall x \forall y((x\neq 0)\wedge(y\neq 0)\leftrightarrow(xy\neq 0))$

25. 将下列嵌套量化式子翻译成表达出数学事实的汉语语句。论域均为全体实数。

a) $\exists x \forall y(xy=y)$

b) $\forall x \forall y(((x<0)\wedge(y<0))\rightarrow(xy>0))$

c) $\exists x \exists y((x^2>y)\wedge(x<y))$

d) $\forall x \forall y \exists z(x+y=z)$

26. 令 $Q(x, y)$ 为语句"$x+y=x-y$"。如果两个变量的论域都是整数集合，下列各项的真值是什么？

a) $Q(1, 1)$

b) $Q(2, 0)$

c) $\forall y Q(1, y)$

d) $\exists x Q(x, 2)$

e) $\exists x \exists y Q(x, y)$

f) $\forall x \exists y Q(x, y)$

g) $\exists y \forall x Q(x, y)$

h) $\forall y \exists x Q(x, y)$

i) $\forall x \forall y Q(x, y)$

27. 假定所有变量的论域都是整数集合，确定下列语句的真值。

a) $\forall n \exists m(n^2<m)$

b) $\exists n \forall m(n<m^2)$

c) $\forall n \exists m(n+m=0)$

d) $\exists n \forall m(nm=m)$

e) $\exists n \exists m(n^2+m^2=5)$

f) $\exists n \exists m(n^2+m^2=6)$

g) $\exists n \exists m(n+m=4 \wedge n-m=1)$

h) $\exists n \exists m(n+m=4 \wedge n-m=2)$

i) $\forall n \forall m \exists p(p=(m+n)/2)$

28. 假定每个变量的论域都是实数集合，确定下列语句的真值。

a) $\forall x \exists y(x^2=y)$

b) $\forall x \exists y(x=y^2)$

c) $\exists x \forall y(xy=0)$

d) $\exists x \exists y(x+y\neq y+x)$

e) $\forall x(x\neq 0\rightarrow \exists y(xy=1))$

f) $\exists x \forall y(y\neq 0\rightarrow xy=1)$

g) $\forall x \exists y(x+y=1)$

h) $\exists x \exists y(x+2y=2 \wedge 2x+4y=5)$

i) $\forall x \exists y(x+y=2 \wedge 2x-y=1)$

j) $\forall x \forall y \exists z(z=(x+y)/2)$

29. 假定命题函数 $P(x, y)$ 的论域由 x 和 y 的序偶组成，其中 x 是 1、2 或 3，y 是 1、2 或 3。用析取和合取写出下列命题。

a) $\forall x \forall y P(x, y)$

b) $\exists x \exists y P(x, y)$

c) $\exists x \forall y P(x, y)$

d) $\forall y \exists x P(x, y)$

30. 重写下列语句，使否定只出现在谓词中(即否定词不在量词外边，也不在含逻辑联接词的表达式外边)。

a) $\neg \exists y \exists x P(x, y)$

b) $\neg \forall x \exists y P(x, y)$

c) $\neg \exists y(Q(y) \wedge \forall x \neg R(x, y))$

d) $\neg \exists y(\exists x R(x, y) \vee \forall x S(x, y))$

e) $\neg \exists y(\forall x \exists z T(x, y, z) \vee \exists x \forall z U(x, y, z))$

31. 表达出下列语句的否定，并且所有的否定词紧跟在谓词之前。

a) $\forall x \exists y \forall z T(x, y, z)$

b) $\forall x \exists y P(x, y) \vee \forall x \exists y Q(x, y)$

c) $\forall x \exists y(P(x, y) \wedge \exists z R(x, y, z))$

d) $\forall x \exists y(P(x, y) \rightarrow Q(x, y))$

32. 表达出下列语句的否定，并且所有的否定词紧跟在谓词之前。

a) $\exists z \forall y \forall x T(x, y, z)$

b) $\exists x \exists y P(x, y) \wedge \forall x \forall y Q(x, y)$

c) $\exists x \exists y(Q(x, y) \leftrightarrow Q(y, x))$　　d) $\forall y \exists x \exists z(T(x, y, z) \vee Q(x, y))$

33. 重写下列语句，使否定只出现在谓词中(即否定词不在量词外边，也不在含逻辑联接词的表达式外边)。

a) $\neg \forall x \forall y P(x, y)$　　b) $\neg \forall y \exists x P(x, y)$

c) $\neg \forall y \forall x(P(x, y) \vee Q(x, y))$　　d) $\neg(\exists x \exists y \neg P(x, y) \wedge \forall x \forall y Q(x, y))$

e) $\neg \forall x(\exists y \forall z P(x, y, z) \wedge \exists z \forall y P(x, y, z))$

34. 用量词表达下列语句，然后取该语句的否定并使否定词不在量词的左边。再用简单语句表达这否定(不要简单地表达为"不是……")。

a)没有人因玩彩票损失超过 1000 美元。　　b)班上有个学生只与另外一个学生在网上交谈过。

c)班上没人恰给另外两个同班同学发过电子邮件。　　d)某个学生已完成本书每道练习。

e)没有学生做过本书每节至少一道练习。

35. 用量词表达下列语句，然后取该语句的否定并使否定词不在量词的左边。再用简单语句表达这否定(不要简单地表达为"不是……")。

a)班上每个学生都恰好选修过本校两门数学课。

b)有人去过世界上除利比亚以外的每个国家。

c)没有人攀登过喜马拉雅山的每座山峰。

d)每位电影演员都或与 Kevin Bacon 同拍过电影，或跟与 Kevin Bacon 同拍过电影的人同拍过电影。

36. 用量词和语句表达下列命题的否定。

a)班上每个学生都喜欢数学。　　b)班上有个学生没见过计算机。

c)班上有个学生选修过本校开设的每门数学课。d)班上有个学生去过校园每座楼的至少一个房间。

37. 找出下列全称量化语句的反例(如果可能的话)，其中所有变量的论域是全体整数。

a) $\forall x \forall y(x^2 = y^2 \rightarrow x = y)$　　b) $\forall x \exists y(y^2 = x)$　　c) $\forall x \forall y(xy \geqslant x)$

38. 找出下列全称量化语句的反例(如果可能的话)，其中所有变量的论域是全体整数。

a) $\forall x \exists y(x = 1/y)$　　b) $\forall x \exists y(y^2 - x < 100)$　　c) $\forall x \forall y(x^2 \neq y^3)$

39. 用量词表达实数乘法的结合律。

40. 用量词表达实数乘法和加法的分配律。

41. 如果变量的论域分别如下，确定语句 $\forall x \exists y(xy = 1)$的真值。

a)非零实数　　b)非零整数　　c)正实数

42. 如果变量的论域分别如下，确定语句 $\exists x \forall y(x \leqslant y^2)$的真值。

a)正实数　　b)整数　　c)非零实数

43. 证明语句$\neg \exists x \forall y P(x, y)$和$\forall x \exists y \neg P(x, y)$有同样的真值。

*44. 证明 $\forall x P(x) \vee \forall x Q(x)$和$\forall x \forall y(P(x) \vee Q(y))$逻辑等价。(新变量 y 用来正确地把量词组合在一起。)

*45. a)证明 $\forall x P(x) \wedge \exists x Q(x)$和$\forall x \exists y(P(x) \wedge Q(y))$逻辑等价。

b)证明 $\forall x P(x) \vee \exists x Q(x)$和$\forall x \exists y(P(x) \vee Q(y))$逻辑等价。

一个语句称为前束范式(PNF)的充分必要条件是其表达形式为

$$Q_1 x_1 Q_2 x_2 \cdots Q_k x_k P(x_1, x_2, \cdots, x_k)$$

其中每个 Q_i，$i = 1, 2, \cdots, k$，或是全称量词，或是存在量词；而 $P(x_1, x_2, \cdots, x_k)$是不含量词的谓词。例如$\exists x \forall y(P(x, y) \wedge Q(y))$是前束范式，而$\exists x P(x) \vee \forall x Q(x)$不是(因为并不是所有量词都先出现)。

将命题变量、谓词、T 和 F 用逻辑联接词和量词组合在一起形成的每一个语句都等价于一个前束范式。练习 47 要求的就是对这一事实的证明。

*46. 把下列语句改为前束范式。(提示：利用 1.2 节表 1-13 和表 1-14 的等价关系，1.3 节表 1-17，1.3 节练习 42～45 和本节的练习 44 和 45。)

a) $\exists x P(x) \vee \exists x Q(x) \vee A$，其中 A 是不含量词的命题。

b) $\neg(\forall x P(x) \vee \forall x Q(x))$

c) $\exists x P(x) \rightarrow \exists x Q(x)$

**47. 给出并证明把任意语句变换为等价的前束范式的方法。

48. 如果 x 大于或等于 S 中的每个数，实数 x 称为实数集合 S 的上界。如果 x 是 S 的上界，而且小于或等于 S 的所有上界，实数 x 称为实数集合 S 的最小上界。如果集合 S 有最小上界，这最小上界一定是唯一的。

a)用量词表达 x 是 S 的上界这一事实。

b)用量词表达 x 是 S 的最小上界这一事实。

*49. 用全称量词、存在量词和逻辑联接词表达量化表达式 $\exists! xP(x)$。

语句 $\lim_{n\to\infty} a_n = L$ 表示对每个正实数 ϵ，有个正整数 N，使得只要 $n>N$，就有 $|a_n - L| < \epsilon$。

50. (需要微积分知识)用量词表示语句 $\lim_{n\to\infty} a_n = L$。

51. (需要微积分知识)用量词表示语句 $\lim_{n\to\infty} a_n$ 不存在。

52. (需要微积分知识)用量词表达下面的定义：如果对每个实数 $\epsilon>0$，存在正整数 N，使得每一对正整数 m 和 n，只要 $m>N$ 和 $n>N$，就有 $|a_m - a_n| < \epsilon$，则序列 $\{a_n\}$ 称为哥西序列。

53. (需要微积分知识)用量词和逻辑联接词表达下面的定义：如果对每个实数 $\epsilon>0$，有无穷多个 n 使 $a_n > L-\epsilon$ 成立，而只有有限多个 n 使 $a_n > L+\epsilon$ 成立，则实数 L 称为序列 $\{a_n\}$ 的最大限。

1.5 证明方法

1.5.1 引言

在数学研究中提出的两个重要问题是：(1)什么时候数学论证是正确的？(2)什么方法可以用来构造数学论证？本节通过描述各种形式的正确与不正确的数学论证来帮助回答这些问题。

定理是可以被证明为真的命题。(定理有时称为命题、事实或结果。)用一系列命题来证明一条定理为真，这些命题就形成一项论证，称为证明。在证明里用到的命题可以包括：公理或公设(它们是关于数学结构的基本假设)，被证明定理本身的前提，以及从前证明过的定理等。推理规则(它们是从其他断言得出结论所用的方法)把证明的各个步骤联系起来。

本节将讨论推理规则。这样做有助于解释清楚正确的证明是由什么组成的。还将描述某些常见形式的不正确的推理(称为谬误)。然后将介绍各种常用的证明定理的方法。

名词"引理"和"推论"用于特定类型的定理。**引理**是在其他定理的证明中所用的简单定理。当使用了一系列的引理(其中每个引理都被单独地证明)时，一些复杂的证明通常会更容易被理解。**推论**是从已经证明了的定理直接证实的命题。"**猜想**"是真值未知的命题。当发现了猜想的证明时，这个猜想就成为定理。许多时候猜想都证明为假，所以猜想不是定理。

本章讨论的证明方法是重要的，不仅因为它们被用来证明数学定理，而且因为它们对计算机科学的诸多应用。这些应用包括：验证计算机程序是正确的，证明操作系统是安全的，在人工智能领域里进行推理，等等。所以，理解证明中所用到的技术无论在数学里还是在计算机科学里都是必要的。

1.5.2 推理规则

现在将要介绍命题逻辑的推理规则。这些规则为用来证明一个结论是从一组前提合乎逻辑地得出的步骤提供了正当理由。重言式 $(p\wedge(p\to q))\to q$ 是一个被称为假言推理或分离规则的推理规则的基础。这个重言式写成下列方式：

$$\begin{array}{l} \quad p \\ \quad \underline{p \to q} \\ \therefore q \end{array}$$

使用这种记号时，前提写成一列，而结论写在横线下。（符号∴表示“因此”。）假言推理是说，若蕴含式及其前提都已知为真，则这个蕴含式的结论为真。

Extra Examples **例 1** 假定蕴含式“若今天下雪，则将去滑雪”和它的前件“今天正在下雪”都为真。那么，根据假言推理，得出蕴含式的后件“将去滑雪”为真。■

例 2 蕴含式“若 n 被 3 整除，则 n^2 被 9 整除”为真。所以，若 n 被 3 整除，则根据假言推理得出 n^2 被 9 整除。■

表 1-19 列出了重要的推理规则。在 1.2 节的练习里可以找到对这些推理规则的验证。这里是使用这些推理规则的一些例子。

表 1-19 推理规则

推理规则	重言式	名称
p $\therefore p \vee q$	$p \to (p \vee q)$	附加
$p \wedge q$ $\therefore p$	$(p \wedge q) \to p$	化简
p q $\therefore p \wedge q$	$((p) \wedge (q)) \to (p \wedge q)$	合取
p $p \to q$ $\therefore q$	$[p \wedge (p \to q)] \to q$	假言推理
$\neg q$ $p \to q$ $\therefore \neg p$	$[\neg q \wedge (p \to q)] \to \neg p$	取拒式
$p \to q$ $q \to r$ $\therefore p \to r$	$[(p \to q) \wedge (q \to r)] \to (p \to r)$	假言三段论
$p \vee q$ $\neg p$ $\therefore q$	$[(p \vee q) \wedge \neg p] \to q$	析取三段论
$p \vee q$ $\neg p \vee r$ $\therefore q \vee r$	$[(p \vee q) \wedge (\neg p \vee r)] \to (q \vee r)$	消解

例 3 说出哪个推理规则是下列论证的基础：“现在气温在冰点以下。因此，要么现在气温在冰点以下，要么现在下雨。”

解 设 p 是命题“现在气温在冰点以下”，而 q 是命题“现在正在下雨”。那么这个论证形如

$$\begin{array}{l} \quad \underline{p \qquad} \\ \therefore p \vee q \end{array}$$

这是使用附加规则的论证。■

例 4 说出哪个推理规则是下列论证的基础：“现在气温在冰点以下并且现在下雨。因此，现在气温在冰点以下。”

解 设 p 是命题"现在气温在冰点以下"，而 q 是命题"现在正在下雨"。这个论证形如

$$\begin{array}{l} \underline{p \wedge q} \\ \therefore p \end{array}$$

这个论证使用化简规则。 ■

例 5 说出在下列论证里使用哪个推理规则：

若今天下雨，则我们今天将不野餐。若我们今天不野餐，则我们明天将野餐。因此，若今天下雨，则我们明天将野餐。

解 设 p 是命题"今天下雨"，设 q 是命题"我们今天将不野餐"，而设 r 是命题"我们明天将野餐"。则这个论证形如

$$\begin{array}{l} p \rightarrow q \\ \underline{q \rightarrow r} \\ \therefore p \rightarrow r \end{array}$$

因此，这个论证是假言三段论。 ■

1.5.3 有效论证

若每当所有的前提都为真时，结论也为真，则这样的论证称为*有效的*。所以，证明从前提 p_1，p_2，…，p_n 合乎逻辑地得出 q，就等于证明蕴含式

$$(p_1 \wedge p_2 \wedge \cdots \wedge p_n) \rightarrow q$$

为真。当在有效论证里用到的所有命题都为真时，就得出正确的结论。不过，当在有效论证里用到一个或多个假命题时，该论证可能得出不正确的结论。例如，"若$\sqrt{2} > \frac{1}{2}$，则$(\sqrt{2})^2 > \left(\frac{3}{2}\right)^2$。由此可知$\sqrt{2} > \frac{3}{2}$。因此，$(\sqrt{2})^2 = 2 > \left(\frac{3}{2}\right)^2 = \frac{9}{4}$。"是一个基于假言推理的有效推理。不过，这个推理的结论为假，因为 9 不能整除 $101^2 = 10201$。在这个论证里使用了假命题"101 被 3 整除"，这就意味着该论证的结论可能为假。

当存在许多前提时，为了证明一个论证是有效的，就常常需要多个推理规则。对此通过下面这些例子来说明。在这些例子中一步一步地显示出论证的步骤，明确地叙述出每步的理由。这些例子也说明如何使用推理规则来分析以用语言表述的论证。

例 6 证明：前提"今天下午没有出太阳并且今天比昨天冷"，"只有今天下午出太阳，我们才将去游泳"，"若我们不去游泳，则我们将乘独木舟游览"，以及"若我们乘独木舟游览，则我们将在黄昏时回家"，导致结论"我们将在黄昏时回家"。

解 设 p 是命题"今天下午出太阳"，q 是命题"今天比昨天冷"，r 是命题"我们将去游泳"，s 是命题"我们将乘独木舟游览"，而 t 是命题"我们将在黄昏时回家"。则这些前提成为$\neg p \wedge q$，$r \rightarrow p$，$\neg r \rightarrow s$，$s \rightarrow t$。结论是 t。

如下构造一个论证来证明这些前提得出需要的结论。

步骤	理由
1. $\neg p \wedge q$	前提引入
2. $\neg p$	化简，用步骤 1
3. $r \rightarrow p$	前提引入
4. $\neg r$	取拒式，用步骤 2 和 3

5. $\neg r \rightarrow s$　　前提引入

6. s　　假言推理，用步骤 4 和 5

7. $s \rightarrow t$　　前提引入

8. t　　假言推理，用步骤 6 和 7 ■

例 7 证明：前提“若你发给我电子邮件消息，则我将完成编写程序”，“若你不发给我电子邮件消息，则我将早早地去睡觉”，以及“若我早早地去睡觉，则我将感觉精力充沛地醒来”，导致结论“若我不完成编写程序，则我将感觉精力充沛地醒来”。

解 设 p 是命题“你发给我电子邮件消息”，q 是命题“我将完成编写程序”，r 是命题“我将早早地去睡觉”，而 s 是命题“我将感觉精力充沛地醒来”。则这些前提是 $p \rightarrow q$，$\neg p \rightarrow r$，$r \rightarrow s$。需要的结论是 $\neg p \rightarrow s$。

下列论证证明这些前提得出需要的结论。

步骤	理由
1. $p \rightarrow q$	前提引入
2. $\neg q \rightarrow \neg p$	步骤 1 和逆否命题
3. $\neg p \rightarrow r$	前提引入
4. $\neg q \rightarrow r$	假言三段论，用步骤 2 和 3
5. $r \rightarrow s$	前提引入
6. $\neg q \rightarrow s$	假言三段论，用步骤 4 和 5 ■

1.5.4 消解

已经开发出自动执行推理和证明定理任务的计算机程序。许多这类程序利用称为消解(resolution)的推理规则。这个推理规则基于永真式：

$$((p \vee q) \wedge (\neg p \vee r)) \rightarrow (q \vee r)$$

(此永真式的验证见 1.2 节练习 28。)消解规则最后的析取 $q \vee r$ 称为消解式(resolvent)。当在此永真式中令 $q=r$ 时，可得$(p \vee q) \wedge (\neg p \vee q) \rightarrow q$。而且，当令 $r=\mathbf{F}$ 时，可得$(p \vee q) \wedge (\neg p) \rightarrow q$(因为 $q \vee \mathbf{F} \Leftrightarrow q$)，这是永真式，析取三段论规则便基于此式。

Extra Examples

例 8 使用消解证明假设“Jasmine 在滑雪或没有下雪”和“下雪了或 Bart 在打曲棍球”蕴含结论“Jasmine 在滑雪或 Bart 在打曲棍球。”

解 令 p 为命题“下雪了”，q 为命题“Jasmine 在滑雪”，r 为命题“Bart 在打曲棍球”。我们可以将假设分别表示为$\neg p \vee q$ 和 $p \vee r$。使用消解，命题 $q \vee r$ 成立。■

消解在基于逻辑规则的编程语言中扮演着重要的角色，如在 Prolog 中(其中用到了量化语句的消解规则)。而且，可以用消解来构建自动定理证明系统。要使用消解来构造命题逻辑中的证明，假设和结论必须被表示为子句，子句是变量的析取或这些变量的否定。我们可以将命题逻辑中非子句的语句用一个或多个等价的是子句的语句替换。例如，假定有一个形如 $p \vee (q \wedge r)$的语句。因为 $p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$，便可以用两个子句 $p \vee q$ 和 $p \vee r$ 来代替 $p \vee (q \wedge r)$。我们可以用语句$\neg p$ 和$\neg q$ 来代替形如$\neg(p \vee q)$的语句，因为德摩根律表明$\neg(p \vee q) \Leftrightarrow \neg p \wedge \neg q$。也可以用等价的析取$\neg p \vee q$ 来代替蕴含 $p \rightarrow q$。

例 9 证明假设$(p \wedge q) \vee r$ 和 $r \rightarrow s$ 蕴含结论 $p \vee s$。

解 可以将假设$(p \wedge q) \vee r$ 重写为两个子句 $p \vee r$ 和 $q \vee r$。也可以将 $r \rightarrow s$ 替换为等价的子句

$\neg r \vee s$。使用子句 $p \vee r$ 和 $\neg r \vee s$，通过消解便可得出 $p \vee s$。■

1.5.5 谬误

几种常见的谬误都来源于不正确的论证。这些谬误看上去像是推理规则，但是它们是基于偶然事件而不是重言式。在这里讨论这些谬误，是为了说明在正确与不正确的推理之间的区别。

命题 $[(p \to q) \wedge q] \to p$ 不是重言式，因为当 p 为假而 q 为真时，它为假。不过，存在许多把它当作重言式的不正确论证。这种类型的不正确推理称为*肯定结论谬误*。

例 10 下列论证是否有效？

若你做本书的每一道练习，则你将学习离散数学。你学习过离散数学。

因此，你做过本书的每一道练习。

解 设 p 是命题"你做过本书的每一道练习"。设 q 是命题"你学习过离散数学"。这个论证形如：若 $p \to q$ 并且 q，则 p。这是使用肯定结论谬误的不正确推理的例子。事实上，你可能通过其他某种方式学习离散数学而没有做本书的每一道练习。（你可能通过阅读、听讲座、做本书的一些但不是全部练习等方式来学习离散数学。）■

命题 $[(p \to q) \wedge \neg p] \to \neg p$ 不是重言式，因为当 p 为假而 q 为真时，它为假。许多不正确的论证都不正确地把它当作了推理规则。这种类型的不正确推理称为*否定假设谬误*。

例 11 设 p 和 q 如同例 10 那样。若蕴含式 $p \to q$ 为真，并且 $\neg p$ 为真，则得出 $\neg p$ 为真这是否正确？换句话说，如果假定：若你做本书里每一道练习，则你将学习离散数学，那么假定：若你没有做过本书里每一道练习，则你没有学习离散数学，这是否正确？

解 即使你没有做过本书里每一道练习，你也可能学过离散数学。这个不正确的论证具有形式：$p \to q$ 和 $\neg p$ 蕴含 $\neg q$，这是否定假设谬误的例子。■

1.5.6 带量词命题的推理规则

已经讨论了命题的推理规则。现在将要描述包含量词的命题的一些重要的推理规则。在数学论证里大量地使用这些推理规则，通常都没有明确地指出来。

*全称例示*是这样的推理规则，用它从前提 $\forall x P(x)$ 得出 $P(c)$ 为真，其中 c 是论域里的具体成员。当从命题"所有女人都聪明"得出"丽沙聪明"时，就使用了全称例示，其中丽沙是论域所有女人中的一员。

*全称生成*是这样的推理规则，它说明在对论域里所有元素 c 来说 $P(c)$ 都为真的前提下，$\forall x P(x)$ 为真。当通过从论域里拿出一个任意元素 c 并证明 $P(c)$ 为真来证明 $\forall x P(x)$ 为真时，就使用了全称生成。所选择的元素 c 必须是论域里一个任意的元素，而不是特定的元素。在许多数学证明里隐含地使用全称量词引入，而很少明确地指出来。

*存在例示*是这样的推理规则，在允许从已经 $\exists x P(x)$ 为真，得出在论域里存在一个使得 $P(c)$ 为真的元素 c。在这里不能选择一个任意值的 c，而必须是使得 $P(c)$ 为真的那个 c。通常不知道 c 是什么，仅仅知道它存在。因为它存在，所以可以给它一个名称（c）而继续论证。

*存在生成*是这样的推理规则，用它在已知使 $P(c)$ 为真的一个具体的 c 时，得出 $\exists x P(x)$ 为真。即若知道论域里一个使 $P(c)$ 为真的元素 c，则知道 $\exists x P(x)$ 为真。

这些推理规则总结在表 1-20 里。在下面例子里将要说明如何使用带量词的命题的推理规则之中的一个。

表 1-20 带量词的命题的推理规则，U 是论域

推理规则	名称
$\dfrac{\forall x P(x)}{\therefore P(c)，若\ c \in U}$	全称量词消去
$\dfrac{P(c)}{\therefore \forall x P(x)}$，对任意 $c \in U$	全称量词引入
$\dfrac{\exists x P(x)}{\therefore P(c)，对某个元素\ c \in U}$	存在量词消去
$\dfrac{P(c)，对某个元素\ c \in U}{\therefore \exists x P(x)}$	存在量词引入

例 12 证明前提“在本离散数学课堂的每一个人学过一门计算机课程”和“玛拉是本课堂的学生”蕴含结论“玛拉学过一门计算机课程”。

解 设 $D(x)$ 表示“x 在本离散数学课堂”，并且设 $C(x)$ 表示“x 学过一门计算机课程。”则前提是 $\forall x(D(x) \to C(x))$ 和 D(玛拉)。结论是 C(玛拉)。

下列步骤可以用来从前提证明结论。

步骤	**理由**
1. $\forall x(D(x) \to C(x))$	前提引入
2. D(玛拉) $\to C$(玛拉)	全称例示，用步骤 1
3. D(玛拉)	前提引入
4. C(玛拉)	假言推理，用步骤 2 和 3

■

例 13 证明前提“在这个班上的某个学生没有读过书”和“班上的每个人都通过了第一门考试”蕴含结论“通过考试的某个人没有读过书”。

解 令 $C(x)$ 表示“x 在这个班中”，$B(x)$ 表示“x 读过书了”，$P(x)$ 为“x 通过了第一门考试。”则前提是 $\exists x(C(x) \wedge \neg B(x))$ 和 $\forall x(C(x) \to P(x))$，结论是 $\exists x(P(x) \wedge \neg B(x))$。下列步骤可以用来从前提证明结论。

步骤	**理由**
1. $\exists x(C(x) \wedge \neg B(x))$	前提引入
2. $C(a) \wedge \neg B(a)$	存在例示，用步骤 1
3. $C(a)$	化简，用步骤 2
4. $\forall x(C(x) \to P(x))$	前提引入
5. $C(a) \to P(a)$	全称例示，用步骤 4
6. $P(a)$	假言推理，用步骤 3 和 5
7. $\neg B(a)$	化简，用步骤 2
8. $P(a) \wedge \neg B(a)$	合取，用步骤 6 和 7
9. $\exists x(P(x) \wedge \neg B(x))$	存在生成，用步骤 8

■

注意 数学论证常常包含既使用命题推理规则又使用量词推理规则的步骤。例如，全称例示和假言推理就常常一起使用。当把这些推理规则组合起来时，从前提 $\forall x(P(x) \to Q(x))$ 和 $P(c)$ 就证明结论 $Q(c)$ 为真，其中 c 是论域的成员。

Extra Examples **注意** 数学里许多定理说：对一个具体集合（比如整数集或实数集）里的所有元素来说成立一个性质。虽然这些定理的准确陈述需要包含全称量词，但是数学里的标准约定是省略全称量词。例如，命题“若整数 n 被3整除，则 n^2 被9整除”。其实意味着“对所有整数 n 来说，若整数 n 被3整数，则 n^2 被9整除”同理，命题“若 $x>y$，其中 x 和 y 都是正实数，则 $x^2>y^2$”其实意味着“对所有正实数 x 和 y 来说，若 $x>y$，则 $x^2>y^2$”。另外，当证明这种类型的定理时，常常使用全称生成规则而不明确地指出来。证明的第一步通常涉及选择论域里的一个一般元素。随后的步骤证明这个元素具有所考虑的性质。全称生成蕴含着对论域里所有元素来说定理都成立。

在随后的讨论里将遵循通常的约定而不明确地指出使用了全称量词和全称生成。不过，读者应当总是能够理解何进隐含地应用了这条推理规则。

1.5.7 证明定理的方法

Assessment 证明定理可能会比较困难，现在进一步明确一下构造证明的方法。将要描述如何证明不同类型的命题。因为许多定理都是蕴含式，所以证明蕴含式的技术是重要的。回忆一下：$p\to q$ 为真，除非 p 为真并且 q 为假。注意当证明命题 $p\to q$ 时，只需要证明：若 p 为真则 q 为真；通常并不要证明 q 为真。下面的讨论将给出最常见的证明蕴含式的技术。

Extra Examples **直接证明** 可以通过证明：若 p 为真则 q 也必然为真，来证明蕴含式 $p\to q$。这样就证明 p 为真而 q 为假的组合永远不会出现。这种证明称为*直接证明*。为了完成这样的证明，假定 p 为真，并且使用推理规则和已经证明的定理，来证明 q 也必然为真。

在给出直接证明的例子前，先给出一个定义。

定义1 整数 n 是偶数，如果存在一个整数 k 使得 $n=2k$；整数 n 是奇数，如果存在一个整数 k 使得 $n=2k+1$。（注意一个整数或为偶数或为奇数。）

例14 给出定理“若 n 是奇数，则 n^2 是奇数”的直接证明。

解 假定这个蕴含式的前件为真。即假定 n 是奇数。则 $n=2k+1$，其中 k 是整数。由此得出 $n^2(2k+1)^2=4k^2+4k+1=2(2k^2+2k)+1$。因此，$n^2$ 是奇数（它是一个整数的两倍再加1）。■

Extra Examples **间接证明** 因为蕴含式 $p\to q$ 等价于它的逆否命题 $\neg q\to\neg p$，所以可以通过证明它的逆否命题 $\neg q\to\neg p$ 为真，来证明蕴含式 $p\to q$。通常直接地证明这个相关的蕴含式，但是任何证明技术都可以使用。这种类型的论证称为*间接证明*。

例15 给出定理“若 $3n+2$ 是奇数，则 n 是奇数”的间接证明。

解 假定这个蕴含式的后件为假：即假定 n 是偶数。则对某个整数 k 来说有 $n=2k$。所以 $3n+2=3(2k)+2=6k+2=2(3k+1)$，所以 $3n+2$ 是偶数（因为它是2的倍数）。因为对这个蕴含式后件的否定蕴含着前件为假，所以原来的蕴含式为真。■

空证明和平凡证明 假定蕴含式 $p\to q$ 的前件 p 为假。则蕴含式 $p\to q$ 为真，因为该命题形如 **F→T** 或 **F→F**，所以它为真。因此，若可以证明 p 为假，则可以给出蕴含式 $p\to q$ 的证明，这称为*空证明*。常常用空证明来证明一些定理的特殊情形，这些定理说：对所有正整数来说，一个蕴含式为真[即形如 $\forall nP(n)$ 的定理，其中 $P(n)$ 是命题函数]。这种定理的证明技术将在3.3节讨论。

例16 证明命题 $P(0)$ 为真，其中 $P(n)$ 是命题函数“若 $n>1$，则 $n^2>n$”。

解 注意命题 $P(0)$ 是蕴含式“若 $0>1$，则 $0^2>0$”。因为前提 $0>1$ 为假，所以蕴含式 $P(0)$ 自

动地为真。 ■

注意 蕴含式的后件 $0^2>0$ 为假这个事实，与蕴含式的真值无关，因为前件为假的蕴含式保证为真。

假定蕴含式 $p\to q$ 的后件 q 为真。则 $p\to q$ 为真，因为该命题形如 $\mathbf{T}\to\mathbf{T}$ 或 $\mathbf{F}\to\mathbf{T}$，而它们都为真。因此，若可以证明 q 为真，则可以给出 $p\to q$ 的证明，这称为平凡证明。当证明定理的特殊情形(见对分情形证明的讨论)时，以及在数学归纳法(它是 3.2 节讨论的一种证明技术)中，平凡证明常常是重要的。

例 17 设 $P(n)$ 是命题“若 a 和 b 是满足 $a\geqslant b$ 的正整数，则 $a^n\geqslant b^n$”。证明命题 $P(0)$ 为真。

解 命题 $P(0)$ 是“若 $a\geqslant b$，则 $a^0\geqslant b^0$”。因为 $a^0=b^0=1$，所以 $P(0)$ 的后件为真。因此，$P(0)$ 为真。这是平凡证明法的一个例子。注意在这个证明里不需要前件，它是命题“$a\geqslant b$。” ■

证明的小策略 我们已经介绍了直接证明和间接证明并给出示例。然而，当面临证明的情形时，要使用哪一种方法呢？首先，快速评估一下直接证明是否可行。这通过扩展前件中的定义开始，然后使用它们(以及公理和可用的定理)进行推理。如果直接证明看来没什么希望，做同样的步骤试一下间接证明。有时当直接证明不太可行时，间接证明却能工作得很好。例 18 和例 19 演示了这种策略。

Extra Examples 在给出示例前，先介绍一个有用的定义。

定义 2 实数 r 是有理数，若存在整数 p 和 $q(q\neq 0)$ 使得 $r=p/q$。不是有理数的实数称为无理数。

例 18 证明两个有理数的和是有理数。

解 首先尝试直接证明。假设 r 和 s 是有理数。由有理数的定义，可知存在整数 p 和 $q(q\neq 0)$ 使得 $r=p/q$，存在整数 t 和 $u(u\neq 0)$ 使得 $s=t/u$。我们能用这个信息证明 $r+s$ 是有理数吗？很显然，下一步是相加 $r=p/q$ 和 $s=t/u$，得到

$$r+s=\frac{p}{q}+\frac{t}{u}=\frac{pu+qt}{qu}$$

因为 $q\neq 0$ 且 $u\neq 0$，所以 $qu\neq 0$。因此，$r+s$ 已经被表示为两个整数 $pu+qt$ 和 qu 的比，其中 $qu\neq 0$。这表示 $r+s$ 是有理数。直接证明成功了。 ■

例 19 证明如果 n 是整数且 n^2 是奇数，则 n 是奇数。

解 首先尝试直接证明。假设 n 是整数且 n^2 是奇数。那么，存在整数 k 使得 $n^2=2k+1$。我们能用这个信息证明 n 是奇数吗？看来没有明显可行的方法，因为求解 n 会得出等式 $n=\pm\sqrt{2k+1}$，这没有什么用。

因为直接证明没有达到目的，下面尝试间接证明。我们将命题“n 不是奇数”作为前件。因为每个整数不是奇数便是偶数，这意味着 n 为偶数。这表示存在整数 k 使得 $n=2k$。要证明这个定理，我们需证明这个前件蕴含着“n^2 不是奇数”的结论，即 n^2 是偶数。我们能用 $n=2k$ 实现这个目标吗？对这个等式的两边取平方，可得 $n^2=4k^2=2(2k^2)$，这表示 n^2 也是偶数，因为 $n^2=2t$ (其中 $t=2k^2$)。间接证明成功了。 ■

归谬证明 当直接证明和间接证明均不成功时，还可以用其他的办法。下面再介绍几种证明技术。

假定可以找到矛盾式 q 使得 $\neg p\to q$ 为真，即 $\neg p\to\mathbf{F}$ 为真。于是命题 $\neg p$ 必然为假。所以 p 必

然为真。当可以找到矛盾式(比如 $r \wedge \neg r$)使得有可能证明蕴含式 $\neg p \to (r \wedge \neg r)$ 为真时，就可以使用这种技术。这种类型的论证称为*归谬证明*。

下面给出3个归谬证明的例子。第一个例子是鸽巢原理(将在4.2节深入介绍)的应用。

Extra Examples **例20** 证明在任意22天中至少有4天属于一个星期的同一天。

解 令 p 为命题“在任意22天中至少有4天属于一个星期的同一天”。假设 $\neg p$ 为真。则22天中至多有3天属于一个星期的同一天。因为一个星期有7天，所以至多可以选择21天。这就与题设矛盾，证毕。 ■

例21 通过给出归谬证明来证明 $\sqrt{2}$ 是无理数。

解 设 p 是命题“$\sqrt{2}$ 是无理数”。假定 $\neg p$ 为真。则 $\sqrt{2}$ 是有理数。将要证明它导致矛盾。在 $\sqrt{2}$ 是有理数的假设下，存在整数 a 和 b 满足 $\sqrt{2}=a/b$，其中 a 和 b 没有公因子(所以分数 a/b 是既约的)。因为 $\sqrt{2}=a/b$，所以当这个等式的两端都平方时，就得出 $2=a^2/b^2$。因此，$2b^2=a^2$。这意味着 a^2 是偶数，它蕴含着 a 是偶数。另外，因为 a 是偶数，所以对某个整数 c 来说有 $a=2c$。因此，$2b^2=4c^2$，所以 $b^2=2c^2$。这意味着 b^2 是偶数。因此，b 也必然是偶数。

已经证明了 $\neg p$ 蕴含着 $\sqrt{2}=a/b$，其中 a 和 b 没有公因子，以及2整除 a 和 b。这是矛盾的，因为已经证明了 $\neg p$ 既蕴含 r 又蕴含 $\neg r$，其中 r 是命题：a 和 b 是没有公因子的整数。因此，$\neg p$ 为假，所以 p：“$\sqrt{2}$ 是无理数”为真。 ■

对一个蕴含式的间接证明可以改写成归谬证明。在间接证明里，通过用直接证明来证明 $\neg q \to \neg p$ 为真，来证明 $p \to q$ 为真。即在 $p \to q$ 的间接证明里，假定 $\neg q$ 为真而证明 $\neg p$ 也必然为真。为了把 $p \to q$ 的间接证明改写成归谬证明，假定 p 和 $\neg q$ 都为真。然后利用 $\neg q \to \neg p$ 的间接证明的步骤，来证明 $\neg p$ 也必然为真。这样就得出矛盾式 $p \wedge \neg p$，由此完成归谬证明。例22说明如何把对蕴含式的间接证明改写成归谬证明。

例22 给出定理“若 $3n+2$ 是奇数，则 n 是奇数”的归谬证明。

解 假定 $3n+2$ 是奇数而 n 不是奇数，所以 n 是偶数。按照在例15解答里的同样步骤(这个定理的间接证明)，可以证明若 n 是偶数则 $3n+2$ 是偶数。这与 $3n+2$ 是奇数的假定矛盾，证毕。 ■

分情形证明 为了证明形如

$$(p_1 \vee p_2 \vee \cdots \vee p_n) \to q$$

的蕴含式，可以用重言式

$$[(p_1 \vee p_2 \vee \cdots \vee p_n) \to q] \leftrightarrow [(p_1 \to q) \wedge (p_2 \to q) \wedge \cdots \wedge (p_n \to q)]$$

Extra Examples 来作为推理规则。这个推理规则说明，可以通过对 $i=1, 2, \cdots, n$ 分别证明每个蕴含式 $p_i \to q$ 来证明由命题 $p_1, p_2, \cdots, p_n$ 的析取式组成前件的原来的蕴含式。这种论证称为*分情形证明*。有时为了证明蕴含式 $p \to q$ 为真，方便的做法是用析取式 $p_1 \vee p_2 \vee \cdots \vee p_n$ 代替 p 作为蕴含式的前件，其中 p 与 $p_1 \vee p_2 \vee \cdots \vee p_n$ 等价。考虑下面的例子。

例23 用分情形证明法证明 $|xy|=|x||y|$，其中 x 和 y 是实数。($|x|$ 是 x 的绝对值，当 $x \geqslant 0$ 时，$|x|=x$；当 $x \leqslant 0$ 时，$|x|=-x$。)

解 令 p 为“x 和 y 是实数”，令 q 为“$|xy|=|x||y|$”。注意 p 等价于 $P_1 \vee P_2 \vee P_3 \vee P_4$，其中 P_1 为“$x \geqslant 0 \wedge y \geqslant 0$”，$P_2$ 为“$x \geqslant 0 \wedge y < 0$”，$P_3$ 为“$x < 0 \wedge y \geqslant 0$”，$P_4$ 为“$x < 0 \wedge y < 0$”。因此，要证明 $p \to q$，我们可以证明 $P_1 \to q$，$P_2 \to q$，$P_3 \to q$ 及 $P_4 \to q$。

很显然 $P_1 \to q$，因为当 $x \geqslant 0$ 且 $y \geqslant 0$ 时 $xy \geqslant 0$，因此 $|xy| = xy = |x||y|$。

要证明 $P_2 \to q$，注意若 $x \geqslant 0$ 且 $y < 0$，则 $xy \leqslant 0$，因此 $|xy| = -xy = x(-y) = |x||y|$。（因此，因为 $y < 0$，我们有 $|y| = -y$。）

要证明 $P_3 \to q$，可遵循前一种情形的推理过程，只需将 x 和 y 的角色互换。

要证明 $P_4 \to q$，注意当 $x < 0$ 且 $y < 0$ 时，$xy > 0$。因此 $|xy| = xy = (-x)(-y) = |x||y|$。证毕。 ■

等价性证明 为了证明一个定理是双条件的，即形如 $p \leftrightarrow q$ 的命题，其中 p 和 q 都是命题，可以使用重言式

$$(p \leftrightarrow q) \leftrightarrow [(p \to q) \wedge (q \to p)]$$

即如果证明了蕴含式"若 p 则 q"和"若 q 则 p"，那么就可以证明命题"p 当且仅当 q"。

例 24 证明定理"整数 n 是奇数当且仅当 n^2 是奇数"。

Extra Examples

解 这个定理是形如"p 当且仅当 q"，其中 p 是"n 是奇数"而 q 是"n^2 是奇数。"为了证明这个定理，需要证明 $p \to q$ 和 $q \to p$ 都为真。

已经证明了（在例 14 里）$p \to q$ 为真且 $q \to p$ 为真（在例 19 里）。

因为已经证明了 $p \to q$ 和 $q \to p$ 都为真，所以就已经证明了这个定理为真。 ■

有时候定理说几个命题都是等价的。这样的定理说命题 p_1，p_2，p_3，…，p_n 都是等价的。这个定理可以写成

$$p_1 \leftrightarrow p_2 \leftrightarrow \cdots \leftrightarrow p_n$$

它说所有的 n 个命题都具有相同的真值。证明这些命题互相等价的一种方式是使用重言式

$$[p_1 \leftrightarrow p_2 \leftrightarrow \cdots \leftrightarrow p_n] \leftrightarrow [(p_1 \to p_2) \wedge (p_2 \to p_3) \wedge \cdots \wedge (p_n \to p_1)]$$

这个重言式说明，若可以证明蕴含式 $p_1 \to p_2$，$p_2 \to p_3$，…，$p_n \to p_1$ 都为真，则命题 p_1，p_2，p_3，…，p_n 都是等价的。

例 25 证明下列三个命题等价：

p_1：n 是偶数

p_2：$n-1$ 是奇数

p_3：n^2 是偶数

解 为了证明这些命题等价，证明蕴含式 $p_1 \to p_2$，$p_2 \to p_3$ 和 $p_3 \to p_1$ 都为真。

将用直接证明来证明 $p_1 \to p_2$ 为真。假定 n 为偶数，则对某个整数 k 来说有 $n = 2k$。因此，$n-1 = 2k-1 = 2(k-1)+1$。这意味着 $n-1$ 是奇数，因为它形如 $2m+1$，其中 $m = k-1$。

仍用直接证明来证明 $p_2 \to p_3$。现在假定 $n-1$ 是奇数，则对某个整数 k 来说有 $n-1 = 2k+1$。因此，$n = 2k+2$，$n^2 = (2k+2)^2 = 4k^2+8k+4 = 2(2k^2+4k+2)$。这意味着 n^2 是整数 $2k^2+4k+2$ 的两倍，所以 n^2 是偶数。

要证明 $p_3 \to p_1$，用间接证明。也就是，证明如果 n 不是偶数，则 n^2 也不是偶数。这等同于证明如果 n 是奇数，那么 n^2 是奇数，这在例 14 中已证明过了。证毕。 ■

1.5.8 定理与量词

许多定理都叙述成带量词的命题。可以使用各种方法来证明量词化的定理。这里将要描述一些最常见的这类方法。

存在性证明 许多定理都断言存在特定类型的对象。这种类型的定理是形如 $\exists x P(x)$ 的命

题，其中 P 是谓词。对形如 $\exists xP(x)$ 的命题的证明称为*存在性证明*。有多种方式来证明这种类型的定理。有时通过找出一个使得 $P(a)$ 为真的元素 a 来给出 $\exists xP(x)$ 的存在性证明。这样的存在性证明称为*构造性的*。给出*非构造性*的证明也是可能的；即不是找出使 $P(a)$ 为真的元素 a，而是以某种其他方式来证明 $\exists xP(x)$ 为真。给出非构造性证明的一种普通方法是使用归谬证明，证明该存在量词化的否定式蕴含着矛盾。下列例子说明构造性的存在性证明的概念。

Extra Examples **例 26 构造性的存在性证明。**证明存在某个正整数，可以用两种不同的方式将其表示为正整数的立方和。

解 经过大量的计算（如使用计算机搜索）可找到

$$1729 = 10^3 + 9^3 = 12^3 + 1^3$$

因为 1729 满足题设要求[⊖]，所以得证。■

例 27 非构造性的存在性证明。证明存在无理数 x 和 y 使得 x^y 是有理数。

解 由例 21 可知 $\sqrt{2}$ 是无理数。考虑数 $\sqrt{2}^{\sqrt{2}}$。如果它是有理数，那就存在两个无理数 x 和 y 且 x^y 是有理数，即 $x=\sqrt{2}$，$y=\sqrt{2}$。另一方面，如果 $\sqrt{2}^{\sqrt{2}}$ 是无理数，那么可以令 $x=\sqrt{2}^{\sqrt{2}}$ 且 $y=\sqrt{2}$，因此 $x^y=(\sqrt{2}^{\sqrt{2}})^{\sqrt{2}}=\sqrt{2}^{(\sqrt{2}\cdot\sqrt{2})}=\sqrt{2}^2=2$。

这是一个非构造性存在性证明的例子，即我们并没有找出满足条件的无理数，而仅仅是证明必然存在这样的无理数。■

唯一性证明 一些定理断言具有特定性质的元素唯一存在。换句话说，这些定理断言恰有一个元素具有这个性质。要证明这类命题，需要证明有某个元素具有这个性质，且没有其他元素有此性质。唯一性证明的两个部分如下：

存在性：证明存在某个元素 x 具有期望的性质。

唯一性：证明若 $y\neq x$，则 y 不具有期望的性质。

注意 *证明存在某个唯一元素 x 等同于证明语句 $\exists x(P(x)\wedge\forall y(y\neq x\rightarrow\neg P(y)))$。*

例 28 证明每个整数都有一个唯一的加性逆。也就是证明，如果 P 是整数，那么存在一个唯一的整数 q 使得 $p+q=0$。

Extra Examples **解** 如果 P 是整数，当 $q=-p$ 且 q 也是整数时，可得 $p+q=0$。因此，存在一个整数 q 使得 $p+q=0$。

要证明 q 的唯一性，假定有某个整数 r 且 $r\neq q$，使得 $p+r=0$。那么 $p+q=p+r$。从等式两边减去 p，得出 $q=r$，这与假定 $q\neq r$ 矛盾。因此，存在某个唯一的整数 q 满足 $p+q=0$。■

Extra Examples **反例** 在 1.3 节曾提到，如果能找到一个反例，便能证明形如 $\forall xP(x)$ 的语句为假。例 29 说明了如何寻找反例。

例 29 证明语句“每个正整数都是三个整数的平方和”为假。

解 如果能找到一个反例，便能证明此语句为假。要寻找反例，我们试着写出可满足语句条件的连续的正整数。可以发现 $1=0^2+0^2+1^2$，$2=0^2+1^2+1^2$，$3=1^2+1^2+1^2$，$4=0^2+0^2+2^2$，$5=0^2+1^2+2^2$，$6=1^2+1^2+2^2$，但我们无法找到方法将 7 写为三个整数的平方和。要证明 7 是一个反例，请注意可用的不超过 7 的平方数只有 0、1 或 4，而三个项（每项为 0、1 或 4）相加却得不出 7，所以 7 是一个反例。

⊖ 历史注记：英国数学家 G. H. 哈代有一次前往医院看望生病的印度天才拉马努金，哈代提到他乘坐的出租车的编号 1729 是个枯燥的数字。拉马努金说“不，1729 非常有趣，它是可以用两种方式来表示为整数立方和的最小的数。”（参见第 3 章补充练习中关于哈代和拉马努金的简介。）

命题得证。 ■

一种常见的错误是假定一个或多个特例就证明了命题为真。无论存在多少个使 $P(x)$ 为真的特例，全称量词式 $\forall xP(x)$ 仍可能为假。考虑下例。

例 30 语句"每个正整数都是 18 个四次方整数之和"是否为真？也就是，命题 $\forall nP(n)$ 是否是定理，其中 $P(n)$ 是语句"n 可以被写为 18 个四次方整数之和"，论域是全体正整数？

解 要判断 n 是否可写为 18 个四次方整数的和，我们先从最小的正整数开始考察。因为整数的四次方分别是 0，1，16，81，…，如果能从这些数中选择 18 个项后相加得 n，则命题得证。可以证明，从 1 到 78 的所有正整数都可以写成 18 个四次方整数的和(细节留给读者证明)。然而，如果认为这就检查够了，那就会得出错误的结论，因为 79 并不是 18 个四次方整数的和(读者请自行验证)。所以，题设语句不为真。 ■

1.5.9 证明中的错误

在构造数学证明时，有一些错误容易犯。这里简述其中的一些。这当中最常见的错误是算术和基本代数方面的。甚至专门的数学家也会犯这种错误，尤其是在处理复杂的公式时。任何时候进行这样的计算都应当尽可能仔细地检查。

数学证明的每一步都应当是正确的，并且结论要从之前的步骤中逻辑地导出。许多错误是源于引入不是前面前步骤的逻辑推导。下面给出几个例子。

例 31 下面这个著名的假定"证明"(即 $1=2$)错在哪里？

"证明"步骤如下，其中 a 和 b 是两个相等的正整数。

步骤	理由
1. $a=b$	给定的前提
2. $a^2=ab$	步骤 1 两边乘以 a
3. $a^2-b^2=ab-b^2$	步骤 2 两边减去 b^2
4. $(a-b)(a+b)=b(a-b)$	步骤 3 两边分解因式
5. $a+b=b$	步骤 4 两边除以 $a-b$
6. $2b=b$	步骤 5 把 a 替换成 b，因为 $a=b$ 并化简
7. $2=1$	步骤 6 两边除以 b

解 除了步骤 5 两边除以$(a-b)$之外，每个步骤都有效，错误在于 $a-b$ 等于零；一个等式两边用同一个数相除只有在除数不是零时才是有效的。 ■

例 32 下面这个"证明"错在哪里？

"定理"：若 n^2 是正数，则 n 是正数。

"证明"：假定 n^2 是正数。因为蕴含"若 n 是正数，则 n^2 是正数"为真，所以可以得出 n 是正数。

解 令 $P(n)$为"n 是正数"，$Q(n)$为"n^2 是正数"。则前提是 $Q(n)$。命题"如果 n 是正数，则 n^2 是正数"也就是 $\forall n(P(n)\rightarrow Q(n))$。从前提 $Q(n)$和 $\forall n(P(n)\rightarrow Q(n))$不能得出 $P(n)$，因为没有使用有效的推理规则。相反，这是一个断言结论的谬误的示例。一个反例是当 $n=-1$ 时，$n^2=1$为正数，但 n 却是负的。 ■

例 33 下面的"证明"错在哪里？

"定理"：如果 n 不是正数，则 n^2 不是正数。

"证明"：假定 n 不是正数。因为蕴含"如果 n 是正数，则 n^2 是正数"为真，所以可得 n^2 不是

正数。

解 如例 32 那样设定 $P(n)$和 $Q(n)$。则前提是$\neg P(n)$，命题“如果 n 是正数，则 n^2 是正数”即是 $\forall n(P(n)\rightarrow Q(n))$。从前提$\neg P(n)$和 $\forall n(P(n)\rightarrow Q(n))$不能得出$\neg Q(n)$，因为没有使用有效的推理规则。相反，这是一个否定前提的谬误示例。如例 32 那样，$n=-1$ 即为反例。 ■

在分情形证明中，一种常见错误是作出无根据的假设，其中未考虑到所有情形。如下例所示。

例 34 下面的“证明”错在哪里？

“定理”：如果 x 是实数，则 x^2 是正实数。

“证明”：令 p_1 为“x 是正数”，p_2 为“x 是负数”，q 为“x^2 是正数”。要证明 $p_1\rightarrow q$，注意，当 x 是正数时，x^2 为正数，因为两个正数 x 和 x 的积肯定为正。要证明 $p_2\rightarrow q$，注意当 x 是负数时，x^2 是正数，因为 x^2 是两个负数的积。证毕。

解 上面的证明存在的问题是忘了考虑 $x=0$ 的情形。当 $x=0$ 时，$x^2=0$ 非正，因此假设的定理为假。如果 p 是“x 是实数”，那我们可以分三种情形 p_1、p_2 和 p_3 来证明结果，其中 p_1 是“x 是正数”，p_2 是“x 是负数”，p_3 是“$x=0$”，因为 $p\leftrightarrow p_1\vee p_2\vee p_3$。 ■

最后，我们简要讨论一种比较难应付的错误。许多不正确的论证都基于一种称为偷用论题的谬误。当证明的一个或多个步骤基于待证明命题的正确性时，会发生这样的谬误。换句话说，当命题使用自身或等价于自身的命题来进行证明时会产生这种谬误。所以这种谬误也称为*循环论证*。

例 35 下面的论证是否正确？

假定 n^2 是偶数，则存在某个整数 k 使 $n^2=2k$。令存在某个整数 l 使得 $n=2l$。这证明了 n 是偶数。

解 这个论证不正确。证明中没有给出论证说明 n 可以写为 $2l$，l 为某整数。这是一个循环论证，因为这个命题等价于待证的命题(即 n 是偶数)。当然，结果本身是正确的，只是证明方法不对。 ■

在证明中犯错是学习过程的一部分。当你犯了某个错误并被别人发现时，应该仔细分析哪里出了错误并确保不再犯同样的错误。

1.5.10 关于证明的一些评注

已经描述了证明定理的各种方法。注意，在这里没有给出甚至没有提到证明定理的算法。一个深奥的结果是不存在这样的过程。

存在着许多定理，通过直接利用前提和定理里的名词的定义，就容易找出其证明。不过，通常要是不借助于灵活地利用间接证明(比如归谬证明)或一些其他的证明技术，就难以证明一个定理。构造证明是一种艺术，只有通过尝试各种使用途径才能学会它。

在回到证明这个主题前，在本章的余下部分及第 2 章中我们将给出多种证明。在第 3 章中，将介绍一些证明定理和猜想时的方法和策略。在第 3 章中也将引入几种重要的证明技术，包括数学归纳法，这种方法可以用于证明对所有正整数成立的结论。在第 4 章将介绍组合证明的概念。

练习

1. 在下列每个论证里使用了什么推理规则？
   a)爱丽丝主修数学。因此，爱丽丝主修数学或计算机科学。
   b)杰瑞主修数学和计算机科学。因此，杰瑞主修数学。
   c)若天气下雨，则游泳池将关闭。天气下雨。因此，游泳池关闭。

d)若今天下雪，则大学将关闭。今天大学没有关闭。因此，今天没有下雪。

e)若我去游泳，则我将在太阳下停留过久。若我在太阳下停留过久，则我将有晒斑。因此，若我去游泳，则我将有晒斑。

2. 在下列每个论证里使用了什么推理规则？

a)袋鼠生活在澳大利亚并且是有袋动物。因此，袋鼠是有袋动物。

b)今天气温高于100度或污染是有害的。今天户外气温低于100度。因此，污染是有害的。

c)琳达是优秀的游泳者。若琳达是优秀的游泳者，则她可以当救生员。因此，琳达可以当救生员。

d)今年夏天史蒂夫将在计算机公司工作。因此，今年夏天史蒂夫将在计算机公司工作或者在海滩闲逛。

e)若我整夜地做这个家庭作业，则我可以解答所有的习题。若我解答所有的习题，则我将理解这些材料。因此，若我整夜地做这个家庭作业，则我将理解这些材料。

3. 使用推理规则构造一个论证来证明前提“兰迪努力地工作”，“若兰迪努力地工作，则他是一个笨孩子”，和“若兰迪是一个笨孩子，则他将得不到这个工作”蕴含着结论“兰迪将得不到这个工作”。

4. 使用推理规则构造一个论证来证明前提“若天气不下雨或天气不起雾，则航行比赛将举行而且救生表演将进行”，“若航行比赛举行，则将颁发奖品”，而“没有颁发奖品”蕴含着结论“天气下雨”。

5. 在下面的著名论证里使用了什么推理规则？“所有的人都是要死的。苏格拉底是人。因此，苏格拉底是要死的”。

6. 在下面的论证里使用了什么推理规则？“没有人是岛屿。曼哈顿是岛屿。因此，曼哈顿不是人。”

7. 对下列的每组前提，可以得出什么样的相关结论或一组结论？解释从前提获得每个结论所使用的推理规则。

a)“若我在某一天休息，则那天下雨或下雪”。“我在周二休息或在周四休息”。“周二出太阳”。“周四未下雪。”

b)“若我吃了辣的食物，则我做奇怪的梦”。“若当我睡觉时有雷声，则我做奇怪的梦”。“我没有做奇怪的梦。”

c)“我聪明或幸运”。“我不幸运”。“若我幸运，则我将赢得抽奖。”

d)“每个主修计算机科学的人都有个人电脑”。“拉尔夫没有个人电脑”。“安妮有个人电脑。”

e)“对公司有利的就对美国有利”。“对美国有利的就对你有利”。“对公司有利的就是你购买许多东西。”

f)“所有的啮齿类动物都啃它们的食物。”“老鼠是啮齿类动物”。“野兔不啃它们的食物”。“蝙蝠不是啮齿类动物”。

8. 对下列的每组前提，可以得出什么样的相关结论或一组结论？解释从前提获得每个结论所使用的推理规则。

a)“若我打曲棍球，则我第二天感到酸痛”。“若我感到酸痛，则我用水疗”。“我没有用水疗”。

b)“若我工作，则天气晴或半晴”。“我上周一工作或上周五工作”。“周二天气不晴”。“周五天气不是半晴。”

c)“所有的昆虫都有六条腿”。“蜻蜓是昆虫”。“蜘蛛不是六条腿”。“蜘蛛吃蜻蜓”。

d)“每个学生都有因特网账号”。“荷马没有因特网账号”。“马奇有因特网账号”。

e)“所有对健康有益的食物都不好吃。”“豆腐对健康有益”。“你只吃好吃的东西”。“你不吃豆腐”。“汉堡包对健康无益”。

f)“我在做梦或在幻觉中”。“我不在做梦”。“若我在幻觉中，则我看见大象在路上跑”。

9. 对下列每个论证，解释对每个步骤使用了哪条推理规则。

a)“本班学生道格知道如何用Java写程序。知道如何用Java写程序的每个人都可以得到高薪的工作。因此，本班的某个人可以得到高薪的工作。”

b)“本班的某个人喜欢观察鲸鱼。每个喜欢观察鲸鱼的人都关心海洋污染。因此，本班里有某个人关心海洋污染。”

c)“本班的93个学生每人拥有一台个人电脑。拥有个人电脑的每个人都能使用字处理软件。因此，本班

学生泽克能使用字处理软件。”

d)“新泽西的每个人都生活在距离海洋 50 英里之内。新泽西的某个人从来没有见过海洋。因此，生活在距离海洋 50 英里之内的某个人从来没有见过海洋。”

10. 对下列每个论证，解释对每个步骤使用了哪条推理规则。

a)“本班学生琳达拥有红色敞篷汽车。拥有红色敞篷汽车的每个人都至少领到一张超速罚单。因此，本班的某个人领到一张超速罚单。”

b)“五位室友中的每一位——梅丽莎、阿茸、拉尔夫、维妮莎和基绍恩——都学习了离散数学课程。每个学习了离散数学课程的学生都可以学习算法课程。因此，所有的五位室友明年都可以学习算法课程。”

c)“塞雷斯”(John Sayles)制作的所有电影都好看。塞雷斯制作过关于煤矿工人的电影。因此，存在好看的关于煤矿工人的电影。”

d)“本班有人到过法国。到过法国的每个人都访问过卢浮宫。因此，本班有人访问过卢浮宫。”

11. 判断下列论证是否正确并解释原因。

a)班上的所有学生都理解逻辑。Xavier 是这个班上的学生，因此，Xavier 理解逻辑。

b)每个计算机专业的学生都要学离散数学。Natasha 在学离散数学，因此，Natasha 是计算机专业的。

c)所有鹦鹉都喜欢水果。我养的鸟不是鹦鹉，因此，我养的鸟不喜欢水果。

d)每天吃麦片的人都很健康。Linda 不健康，因此，Linda 没有每天吃麦片。

12. 判断下列论证是否正确并解释原因。

a)每一个上大学的人都住过宿舍。Mia 从未住过宿舍，因此，Mia 没有上过大学。

b)敞篷轿车开起来感觉很好。Isaac 的车不是敞篷的，因此，Isaac 的车开起来感觉不好。

c)Quincy 喜欢所有的动作电影。Quincy 喜欢电影《*Eight Men Out*》，因此，《*Eight Men Out*》是动作电影。

d)所有捕虾者都设置了至少 12 个陷阱。Hamilton 是捕虾者，因此，Hamiltion 设置了至少 12 个陷阱。

13. 判定下列每个论证是否有效。若论证是正确的，则使用了什么推理规则？若它不正确，则发生了什么谬误？

a)若 n 是实数使得 $n>1$，则 $n^2>1$。假定 $n^2>1$。于是 $n>1$。

b)若数 $\log_2 3$ 不是两个整数的商，则它是无理数。因此，因为 $\log_2 3$ 不能写成 a/b 的形式，其中 a 和 b 是整数，所以它是无理数。

c)若 n 是使 $n>3$ 的实数，则 $n^2>9$。假定 $n^2\leqslant 9$，于是 $n\leqslant 3$。

d)若 n 是使 $n>2$ 的实数，则 $n^2>4$。假定 $n\leqslant 2$，于是 $n^2\leqslant 4$。

14. 判断下列论证是否有效。

a)如果 x^2 是无理数，则 x 是无理数。因此，如果 x 是无理数，则可知 x^2 是无理数。

b)如果 x^2 是无理数，则 x 是无理数。数 $x=\pi^2$ 是无理数，因此 $x=\pi$ 是无理数。

15. 如下论证错在哪里？令 $H(x)$ 为“x 很开心”。给定前提 $\exists xH(x)$，我们得出 $H(\text{Lola})$。因此，LoLa 很开心。

16. 如下论证错在哪里？令 $S(x, y)$ 为“x 比 y 矮”。给定前提 $\exists sS(s, \text{Max})$，可得出 $S(\text{Max}, \text{Max})$。则由存在量词引入，可得出 $\exists xS(x, x)$，因此某人比他自身矮。

17. 证明命题 $P(0)$，其中 $P(n)$ 是命题“若 n 是个大于 1 的正整数，则 $n^2>n$”。你使用什么类型的证明？

18. 证明命题 $P(1)$，其中 $P(n)$ 是命题“若 n 是个正整数，则 $n^2>n$”。你使用什么类型的证明？

19. 设 $P(n)$ 是命题“若 a 和 b 是正实数，则 $(a+b)^n\geqslant a^n+b^n$”。证明 $P(1)$ 为真。你使用什么类型的证明？

20. 证明：偶数的平方是偶数，使用

a)直接证明。　　b)间接证明。　　c)归谬证明。

21. 证明：若 n 是整数而且 n^3+5 是奇数，则 n 是偶数，使用

a)间接证明。　　b)归谬证明。

22. 证明：若 n 是整数而且 $3n+2$ 是偶数，则 n 是偶数，使用
    a)间接证明。 b)归谬证明。
23. 证明：两个奇数之和是偶数。
24. 证明：两个奇数之积是奇数。
25. 证明：一个无理数与一个有理数之和是无理数，使用归谬证明。
26. 证明：两个有理数之积是有理数。
27. 证明或反驳：两个无理数之积是无理数。
28. 证明或反驳：一个非零有理数与一个无理数之积是无理数。
29. 证明：若 x 是无理数，则 $1/x$ 是无理数。
30. 证明：若 x 是有理数且 $x\neq 0$，则 $1/x$ 是有理数。
31. 证明：在任意 64 天中至少有 10 天指向一个星期的同一天。
32. 证明：在任意 25 天中至少有 3 天指向一年中的同一个月。
33. 证明：若 x 和 y 都是实数，则 $\max(x, y)+\min(x, y)=x+y$。[提示：使用分情形证明，两种情形分别对应于 $x\geqslant y$ 和 $x<y$。]
34. 使用分情形证明来证明：每当 a，b 和 c 都是实数时就有 $\min(a, \min(b, c))=\min(\min(a, b), c)$。
35. 证明：若 x 和 y 都是实数，则 $|x|+|y|\geqslant|x+y|$（其中 $|x|$ 表示 x 的绝对值，若 $x\geqslant 0$ 则它等于 x，若 $x<0$ 则它等于 $-x$）。
36. 证明：任意整数的平方均是以 0，1，4，5，6 或 9 结尾的数。（提示：令 $n=10k+l$，其中 $l=0, 1, \cdots, 9$。）
37. 证明：任意整数的四次方均是以 0，1，5 或 6 结尾的数。
38. 证明：若 n 是正整数，则 n 是偶数当且仅当 $7n+4$ 是偶数。
39. 证明：若 n 是正整数，则 n 是奇数当且仅当 $5n+6$ 是奇数。
40. 证明：$m^2=n^2$ 当且仅当 $m=n$ 或 $m=-n$。
41. 证明或反驳：若 m 和 n 是整数使得 $mn=1$，则 $m=1$ 并且 $n=1$，或者 $m=-1$ 并且 $n=-1$。
42. 证明下面三个命题等价，其中 a 和 b 是实数：(i)a 小于 b，(ii)a 和 b 的平均值大于 a，(iii)a 和 b 的平均值小于 b。
43. 证明下面三个命题等价：(i)$3x+2$ 是偶数，(ii)$x+5$ 是奇数，(iii)x^2 是偶数。
44. 证明下面三个命题等价：(i)x 是有理数，(ii)$x/2$ 是有理数，(iii)$3x-1$ 是有理数。
45. 证明下面三个命题等价：(i)x 是无理数，(ii)$3x+2$ 是无理数，(iii)$x/2$ 是无理数。
46. 下列求解方程 $\sqrt{2x^2-1}=x$ 的推理过程是否正确？(1) $\sqrt{2x^2-1}=x$ 是给定的；(2) $2x^2-1=x^2$，对(1)的两边平方；(3) $x^2-1=0$，从(2)的两边都减去 x^2；(4) $(x-1)(x+1)=0$，对 x^2-1 进行因式分解；(5) $x=1$ 或 $x=-1$，因为 $ab=0$ 蕴含 $a=0$ 或 $b=0$。
47. 下列求解方程 $\sqrt{x+3}=3-x$ 的步骤是否正确？(1) $\sqrt{x+3}=3-x$ 是给定的；(2) $x+3=x^2-6x+9$，对(1)的两边平方；(3) $0=x^2-7x+6$，从(2)的两边都减去 $x+3$；(4) $0=(x-1)(x-6)$，对 x^2-7x+6 进行因式分解；(5) $x=1$ 或 $x=6$，因为 $ab=0$ 蕴含 $a=b$ 或 $b=0$，故从(4)可得到解。
48. 证明存在某个正整数等于小于它的正整数的和。你的证明是构造性的还是非构造性的？
49. 证明存在 100 个连续的不是完全平方的正整数。你的证明是构造性的还是非构造性的？
50. 证明 $2\times 10^{500}+15$ 或 $2\times 10^{500}+16$ 不是完全平方的。你的证明是构造性的还是非构造性的？
51. 证明：存在一对连续的整数，其中一个整数是完全平方的，另一个是完全立方的。
52. 证明：$65^{1000}-8^{2001}+3^{177}$，$79^{1212}-9^{2399}+2^{2001}$ 和 $24^{4493}-5^{8192}+7^{1777}$ 这三个数中任意两个数之积是非负的。你的证明是构造性的还是非构造性的？（提示：不要去求这些数！）
53. 证明下列每一个命题均可用于表达这样的事实：存在一个唯一的元素 x 使得 $P(x)$ 为真。（注意在 1.3 节练习 48 中，这等同于命题 $\exists!\, P(x)$。）

a) $\exists x\forall y(P(y)\leftrightarrow x=y)$　b) $\exists xP(x)\wedge\forall x\forall y(P(x)\wedge P(y)\rightarrow x=y)$

c) $\exists x(P(x)\wedge\forall y(P(y)\rightarrow x=y))$

54. 证明若 a、b 和 c 是实数且 $a\neq 0$，则方程 $ax+b=c$ 存在唯一的解。
55. 假定 a 和 b 是奇数且 $a\neq b$。证明存在唯一的整数 c 满足 $|a-c|=|b-c|$。
56. 证明：若 r 是无理数，则存在唯一的整数 n，使得 r 和 n 之间的距离小于 1/2。
57. 证明：若 n 是奇数，则存在唯一的整数 k，使得 n 是 $k-2$ 和 $k+3$ 之和。
58. 证明：给定实数 x，存在唯一的数 n 和 ε 使得 $x=n+\varepsilon$，n 是整数且 $0\leqslant\varepsilon<1$。
59. 证明：给定实数 x，存在唯一的数 n 和 ε 使得 $x=n-\varepsilon$，n 是整数且 $0\leqslant\varepsilon<1$。
60. 使用消解法证明前提“Allen 是一个坏男孩或 Hillary 是一个好女孩”和“Allen 是一个好男孩或 David 很开心”蕴含结论“Hillary 是一个好女孩或 David 很开心”。
61. 使用消解法证明前提“天上没有下雨或 Yvette 带雨伞了”、“Yvette 没有带雨伞或她没有被淋湿”和“天上下雨了或 Yvette 没有被淋湿”蕴含“Yvette 没有被淋湿”。
62. 用消解法证明可导出等价性 $p\wedge\neg p\equiv\mathbf{F}$。(提示：令 $q=r=\mathbf{F}$。)
63. 用消解法证明复合命题 $(p\vee q)\wedge(\neg p\vee q)\wedge(p\vee\neg q)\wedge(\neg p\vee\neg q)$ 不是可满足的。
64. 证明或反驳：若 a 和 b 是有理数，则 a^b 也是有理数。
65. 证明或反驳：存在有理数 x 和无理数 y，使得 x^y 是无理数。
66. 证明：通过证明 $p_1\leftrightarrow p_4$，$p_2\leftrightarrow p_3$ 和 $p_1\leftrightarrow p_3$，可以证明命题 p_1，p_2，p_3 和 p_4 等价。
67. 证明：可以通过证明蕴含式 $p_1\rightarrow p_4$，$p_3\rightarrow p_1$，$p_4\rightarrow p_2$，$p_2\rightarrow p_5$ 和 $p_5\rightarrow p_3$ 来证明命题 p_1，p_2，p_3，p_4 和 p_5 等价。
68. 证明：用多米诺骨牌(1×2 大小)可以完全覆盖 8×8 的棋盘。

*69. 证明：不可能用多米诺骨牌完全覆盖切掉两个对角格子的 8×8 的棋盘。

*70. 逻辑问题，来自 WFF'N PROOF 中的逻辑游戏，有下面两个假设：

1)“逻辑是困难的或没有许多学生喜欢逻辑”。

2)“若数学是容易的，则逻辑不是困难的”。

把这些假设翻译成包括命题变元和逻辑联接词的命题，判定下面每个命题是不是这些前提的有效结论：

a)若有许多学生喜欢逻辑，则数学是容易的。

b)若数学不是容易的，则没有许多学生喜欢逻辑。

c)数学不是容易的或逻辑是困难的。

d)逻辑不是困难的或数学不是容易的。

e)若没有许多学生喜欢逻辑，则数学不是容易的或逻辑不是困难的。

71. 证明：在实数 a_1，a_2，…，a_n 中至少有一个数大于或等于这些数的平均值。你使用什么类型的证明？
72. 使用练习 71 来证明：若把前 10 个正整数以任意顺序放在一个圆周上，则圆周上存在位置相邻的 3 个整数，它们之和大于或等于 17。
73. 证明：若 n 是整数，则下面 4 个命题等价：(i)n 是偶数；(ii)$n+1$ 是奇数；(iii)$3n+1$ 是奇数；(iv)$3n$ 是偶数。
74. 证明下面 4 个命题等价：(i)n^2 是奇数，(ii)$1-n$ 是偶数，(iii)n^3 是奇数，(iv)n^2+1 是偶数。
75. 哪些推理规则被用来证明在 1.3 节例 19 里所描述的卡洛尔(Lewis Carroll)的论证的结论？
76. 哪些推理规则被用来证明在 1.3 节例 20 里所描述的卡洛尔(Lewis Carroll)的论证的结论？

*77. 判定下列论证(来自 *Back house*[*Ba86*])是否有效：

假如超人能够并愿意防止邪恶，则他将这样做。假如超人不能够防止邪恶，则他将是无能的；假如超人不愿意防止邪恶，则他将是恶意的。超人没有防止邪恶。若超人存在，则他是无能的或恶意的。因此，超人不存在。

1.6 集合

1.6.1 引言

我们将在本书中学习各种各样的离散结构，其中包括由元素的有序偶组成的关系，由元素的无序偶组成的组合，以及由顶点集合和连接顶点的边集合构成的图。不仅如此，我们还将说明这些离散结构以及其他离散结构怎样用于模拟和解题。特别要讨论离散结构用于数据存储、通信和数据处理的许多例子。本节将学习的集合是构造所有其他离散结构的基础。

集合用于把对象组织在一起。通常一个集合中的对象都有相似的性质。例如，你们学校中目前在册的所有学生可以组成一个集合。同样，任何学校正选修一门离散数学课的学生可以组成一个集合。此外，你们学校在册学生中正选修一门离散数学课的所有学生组成一个集合，这个集合可以从上述两个集合中取共有的元素得到。集合语言是以构造的方式研究这种集合的工具。

定义 1 集合是一组无序的对象。

注意，我们并没有说明对象是什么就已经使用这个术语了。以对象的直观概念为基础把集合描述为一组对象，这是德国数学家康托尔㊀1895年首先给出的。从集合的这一直观定义得到的理论导至出悖论，即逻辑上的不一致性。这是英国哲学家罗素㊁1902年提出的(参见练习26，其中描述了一个悖论)。以称为公理的基本假设为起点建立集合理论可以避免逻辑上的不一致。我们将不展开集合理论的公理描述，而是使用康托的称为朴素集合理论的原始描述，这是因为本书中考虑的所有集合均可用康托的原始理论处理而无不一致性。

现在开始讨论集合。

定义 2 集合中的对象也称为该集合的*元素*，或*成员*。也说集合包含它的元素。

有几种方式描述集合，一种方式是在可能的情况下一一列出集合中的元素。我们采用在花括号之间列出所有元素的方法。例如，$\{a, b, c, d\}$表示含4个元素a、b、c和d的集合。

例 1 英语字母中所有元音字母的集合V可以表示为$V=\{a, e, i, o, u\}$。 ■

例 2 小于10的正奇数集合O可以表示为$O=\{1, 3, 5, 7, 9\}$。 ■

㊀ **康托尔(Georg Cantor, 1845—1918)**出生于俄罗斯的圣彼德堡，他父亲是那里一名成功的商人。康托十多岁时对数学产生了浓厚的兴趣。1862年他在苏黎世开始了他的大学学习，不过在他父亲去世时就离开了那里。1863年他在柏林大学继续学习，并得到著名数学家Weierstrass、Kummer和Kronecker的指导。1867年在完成了一篇数论方面的博士论文后他获得博士学位。1869年康托得到哈雷大学的一个职位，并在那里一直工作到去世。

康托是公认的集合论奠基人。他在这一领域的贡献包括实数集合不可数性的发现。他在数学分析方面的贡献也引人注目。康托对哲学也有兴趣，并写了若干论文，将他在集合理论上的工作与形而上学联系在一起。

康托1874年结婚，有5个子女。他忧郁的气质与妻子的乐观性情相互平衡。尽管他从父亲那里得到大笔遗产，但作为教授的收入却很少。为此他曾试图得到柏林大学一个待遇更高的位置。对他的这一任命被Kronecker阻止了，因为Kronecker不同意康托集合论的观点。康托晚年受到精神疾病的折磨，1918年死于精神病诊所。

㊁ **罗素(Bertrand Russell, 1872—1970)**罗素生于一个以积极参与进步运动，热烈地投身自由事业而著名的英格兰家庭。年幼时就成为孤儿的罗素由祖父母扶养，并在家里接受教育。1890年他进入剑桥的Trinity学院学习数学和伦理学。他在几何学基础方面的工作为他赢得了一个研究员职位。1910年Trinity学院任命他教授逻辑和数学原理的课程。

罗素毕生为进步事业而战斗。他有强烈的和平主义观点，他对第一次世界大战的抗议使他失去了Trinity学院的位置。由于一篇被认为具有煽动性的文章使他在1918年被囚禁6个月。罗素还为英国妇女的选举权而斗争。1961年在他89岁高龄时第二次入狱，原因是加入呼吁核裁军的抗议活动。

罗素最伟大的工作是他提出的可以作为所有数学学科基础的原理。他最著名的文章是与怀特海德(Altred North Whitehead)合写的《*Principia Mathematica*》。这篇文章试图用一组基本公理推导出所有数学。他还撰写了许多书籍，内容包括哲学、物理学和他的政治观点。1950年罗素赢得诺贝尔文学奖。

例 3 尽管集合常用来表示一组具有共同性质的元素，但有的集合中的元素表面上看起来毫不相干。例如，集合{ a，2，Fred，New Jerseg }包含了 4 个元素：a、2、Fred 和 New Jerseg。 ■

通常用大写字母表示集合。黑体字母 **N**，**Z**，$\mathbf{Z}^+$，**Q** 和 **R** 分别专用于表示自然数集合{0，1，2，3，…}，整数集合{…，−2，−1，0，1，2，…}，正整数集合，有理数集合$\{p/q \mid p\in \mathbf{Z}, q\in \mathbf{Z}, q\neq 0\}$和实数集合。（有人不认为 0 是自然数，所以在你阅读其他书籍时要细心检查术语自然数的用法。）

有时用花括号表示集合但并不列出它的所有元素。先列出集合中的某些元素，然后当元素的一般形式很明显时就用省略号(…)表示。

例 4 小于 100 的正整数集合可以表示为{1，2，3，…，99}。 ■

由于许多数学语句断定两组以不同方式描述的对象实际上是同一个集合，我们需要理解两个集合相等的含义。

定义 3 两个集合相等当且仅当它们有同样的元素。

例 5 集合{1，3，5}和{3，5，1}相等，因为它们有同样的元素。注意列出元素的顺序不起作用，还要注意同一个元素被列出来不止一次也没关系。所以{1，3，3，3，5，5，5，5}和{1，3，5}是同一个集合，因为它们有同样的元素。 ■

描述集合的另一方式是使用集合构造符号。我们给出作为集合的成员必须具有的性质，以此来刻画集合的所有元素。例如，小于 10 的所有奇数的集合 O 可以写成

$$O = \{x \mid x \text{ 是小于 10 的奇数}\}$$

在无法给出集合所有元素时常用这一类方法描述集合。例如，所有实数的集合可以写成

$$\mathbf{R} = \{x \mid x \text{ 为实数}\}$$

还可以用文氏图形象地表示集合。文氏图是以英国数学家 John Venn⊖ 的名字命名的，他在 1881 年介绍了这种图的使用。我们所考虑的所有对象的集合 U，称为全集。在文氏图中全集用长方形表示。在长方形内部，圆或其他几何图形用于表示集合，有时用点来表示集合中特定的元素。文氏图常用于表示集合之间的关系。我们用下面的例子解释怎样使用文氏图。

例 6 画一个表示英语字母中元音字母集合 V 的文氏图。

解 画一个长方形表示全集 U，也就是 26 个英文字母的集合。在长方形中画一个圆表示集合 V，在圆中用点表示集合 V 的元素(见图 1-1)。 ■

我们现在介绍描述集合成员关系的记号。我们用 $a\in A$ 表示 a 是集合 A 的一个元素。记号 $a\notin A$表示 a 不是集合 A 的成员。注意小写字母通常用于表示集合元素。

有一个不含任何元素的特殊集合，称为空集，用$\varnothing$表示。空集也可以用{}表示(这是用花括号内列出所有元素的方式表示集合)。经常具有一定性质的元素组成的集合其实就是空集。例如，大于自身的平方的所有正整数的集合是空集。

通常的错误是把空集$\varnothing$与**单元集**(即一个元素的集合)集合$\{\varnothing\}$弄错。集合$\{\varnothing\}$的一个元素是集合本身的空集！



⊖ **文氏(John Venn，1834—1923)**出生于伦敦郊区的一个慈善家庭。他在伦敦上学，并于 1857 年获得剑桥 Caius 学院的数学学位。他被选为这一学院的研究员并任此职直至去世。1859 年他接受牧师的圣职，但在短短的宗教工作以后回到了剑桥，并在那里创建了伦理学教育。除数学方面的工作之外，文氏还对历史有兴趣，他写了大量关于他的学院和家庭的作品。

文氏所著《*Symbolic Logic*》一书澄清了最初由布尔引入的若干思想。在这本书中文氏提出了一种系统的研究方法，其中使用了后人称为文氏图的几何图形。今天这些图形主要用于分析逻辑论证及说明集合之间的关系。文氏的成绩不仅是符号逻辑方面的工作，还包括他对概率论的贡献，这些贡献在他编写的广为采用的概率论教科书中都能找到。

定义 4 集合 A 是集合 B 的子集当且仅当 A 的每个元素也是 B 的元素。我们用记号 $A\subseteq B$ 表示 A 是 B 的子集。

我们看到，$A\subseteq B$ 当且仅当量化语句

$$\forall x(x \in A \to x \in B)$$

为真。例如，小于 10 的所有奇数的集合是小于 10 的所有正整数的集合的子集。你们学校主修计算机科学的学生的集合是你们学校全体学生集合的子集。

定理 1 表明，每个非空集合 S 都至少有两个子集，一个空集，一个是集合 S 本身，即 $\varnothing\subseteq S$ 和 $S\subseteq S$。

定理 1 对于任意集合 S，(i) $\varnothing\subseteq S$ 和 (ii) $S\subseteq S$。

证 这里证明(i)，(ii)留作习题。

令 S 为一个集合。为了证明 $\varnothing\subseteq S$，必须证明 $\forall x(x\in \varnothing\to x\in S)$ 为真。因为空集没有元素，所以 $x\in \varnothing$ 总是假。因此 $x\in \varnothing\to x\in S$ 总是真，因为其假设为假，即假的假设为真。即 $\forall x(x\in \varnothing\to x\in S)$ 为真。这完成了(i)的证明。这是空证明的示例。 ■

如果我们要强调集合 A 是集合 B 的子集，但 $A\neq B$，就写成 $A\subset B$，并说 A 是 B 的**真子集**。文氏图可以用来表示集合 A 是集合 B 的子集。我们把全集 U 画成长方形。在这长方形中画一表示 B 的圆。由于 A 是 B 的子集，我们在代表 B 的圆内画表示 A 的圆。图 1-2 所示就是这一关系。

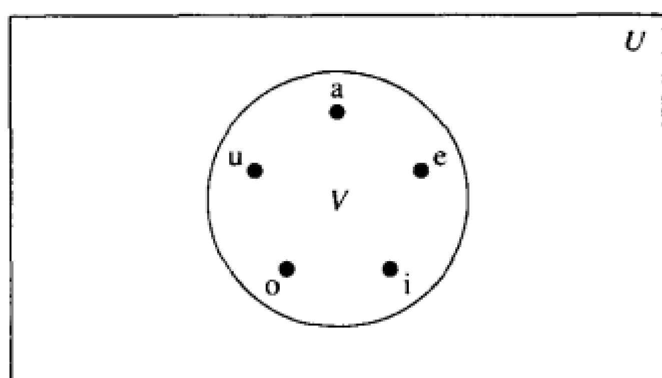


图 1-1 元音字母集合的文氏图

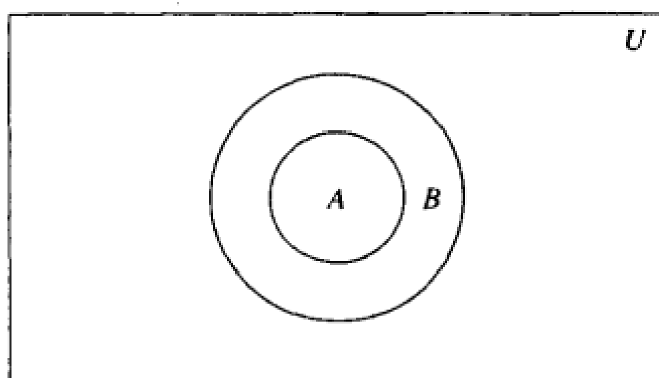


图 1-2 表示 A 是 B 的子集的文氏图

证明两个集合具有相同元素的一个方法是证明它们互为另一个的子集。换言之，可以证明，如果 A 和 B 均为集合，且 $A\subseteq B$ 和 $B\subseteq A$，则 $A=B$。实践表明，这是证明集合相等的一种有用的方式。即 $A=B$，其中 A 和 B 是集合，当且仅当 $\forall x(x\in A\to x\in B)$ 和 $\forall x(x\in B\to x\in A)$；或者 A 和 B 等价，当且仅当 $\forall x(x\in A\leftrightarrow x\in B)$。

集合可以以其他集合作为它的成员。例如，下面列出的就是这样的集合：

$$\{\varnothing,\{a\},\{b\},\{a,b\}\} \text{ 和 } \{x \mid x \text{ 是集合}(a,b) \text{ 的子集}\}$$

注意，这两个集合相等。

集合广泛用于计数问题。这类问题需要讨论集合的大小。

定义 5 令 S 为集合。若 S 中恰有 n 个不同的元素，n 是非负整数，就说 S 是**有限集合**，而 n 是 S 的**基数**。S 的基数用 $|S|$ 表示。

例 7 令 A 为小于 10 的正奇数集合，则 $|A|=5$。 ■

例 8 令 S 为英语字母集，那么 $|S|=26$。 ■

例 9 由于空集没有元素，所以 $|\varnothing|=0$。 ■

我们对不是有限的集合也有兴趣。

定义 6 如果一个集合不是有限的，就说它是**无限的**。

例 10 正整数集合是无限的。 ■


无限集合的基数将在3.2节讨论。在那一节将讨论可数集合的含义，并证明有些集合是可数的，有些不是可数的。

1.6.2 幂集合

许多问题都需要检查一个集合的元素所有可能的组合，看它们是否具有某种性质。为了考虑集合元素所有可能的组合，我们构造一个新集合，它以S的所有子集作为它的元素。

定义7 已知集合S，S的幂集合是集合S所有子集的集合。S的幂集合用$P(S)$表示。

例11 集合$\{0, 1, 2\}$的幂集合是什么？

解 幂集合$P(\{0, 1, 2\})$是$\{0, 1, 2\}$所有子集的集合。因此

$$P(\{0,1,2\}) = \{\varnothing,\{0\},\{1\},\{2\},\{0,1\},\{0,2\},\{1,2\},\{0,1,2\}\}$$

注意，空集和$\{0, 1, 2\}$自身都是这个子集集合的成员。■

例12 空集的幂集合是什么？集合$\{\varnothing\}$的幂集合是什么？

解 空集只有一个子集，这就是它自己。因此

$$P(\varnothing) = \{\varnothing\}$$

集合$\{\varnothing\}$有两个子集，即$\varnothing$和集合$\{\varnothing\}$自身。于是

$$P(\{\varnothing\}) = \{\varnothing,\{\varnothing\}\}$$ ■

如果一个集合有n个元素，那么它的幂集合有2^n个元素。我们将在本书以后的几节中以不同的方式证明这一事实。

1.6.3 笛卡儿积

一组元素中元素的次序往往是重要的。由于集合是无序的，必须用不同的结构来表示有序的一组元素，这个结构由有序n元组来提供。

定义8 有序n元组$(a_1, a_2, \cdots, a_n)$是以a_1为第1个元素，a_2为第2个元素，…，a_n为第n个元素的有序组。

只有在两个有序n元组每一对对应的元素都相等时，它们才相等。换言之，$(a_1, a_2, \cdots, a_n)=(b_1, b_2, \cdots, b_n)$当且仅当对于$i=1, 2, \cdots, n$，$a_i=b_i$。2元组特称为有序偶。有序偶$(a, b)$和$(c, d)$相等当且仅当$a=c$和$b=d$。注意，除非$a=b$，否则$(a, b)$和$(b, a)$不相等。

在随后几章中将要学习的许多离散结构均以René Descartes[⊖]命名的集合的笛卡儿积的概念为基础。我们先定义两个集合的笛卡儿积。


⊖ **笛卡儿(René Descartes, 1596—1650)**出生于法国Tours的一个贵族家庭，Tours位于巴黎西南约200英里处。他是他父亲第一位妻子的第三个孩子，在他出生之后母亲就去世了。由于笛卡儿健康欠佳，他父亲，一位省里的法官，没有让他接受正规教育，直到8岁他才进入La Flèche的Jesuit学院。这所学校的校长喜欢他，并因他身体虚弱而允许他晚起床。从那时起笛卡儿即把早晨时间花在床上，他自己认为这是他思考最富成果的时间。

1612年笛卡儿离开学校去了巴黎，并在那里学了两年数学。1616年他获得Poitiers大学的法律学位。18岁时笛卡儿厌倦了学习，决定看看外部世界。后来他搬到了Saint-Germain城的郊区，一心专注于数学学习。他曾参军，不过他从未参加过战斗。一天，当他为躲避寒冷而呆在军营一间过热的房间里时，他做了几个热昏了的梦，梦中显示他将以做数学家和哲学家为职业。

结束军旅生涯以后，他游遍欧洲。然后在巴黎过了几年，学习数学和哲学，并制造了几个光学仪器。笛卡儿决定去荷兰，并在那里四处漫游20年，完成了他最重要的工作。这期间他写了几本书，包括《*Discours*》，该书中有他对解析几何的贡献，这是他最出名的工作。笛卡儿对哲学也有基础性的贡献。

1649年笛卡儿受克里斯蒂娜女王邀请访问瑞典宫庭，并做她的哲学老师。尽管他不情愿生活在他所说的“岩石和冰雪之间的狗熊的乐土”，但最终还是接受邀请，迁到了瑞典。不幸的是，1649—1650年的冬季分外寒冷，笛卡儿染上了肺炎，于二月中旬去世。

定义 9 令 A 和 B 为集合。A 和 B 的笛卡儿积用 $A\times B$ 表示，是所有有序偶 (a, b) 的集合，其中 $a\in A$ 而 $b\in B$。于是

$$A\times B=\{(a,b)\mid a\in A\wedge b\in B\}$$

例 13 令 A 为某大学所有学生的集合，B 表示该大学开设的所有课程的集合。A 和 B 的笛卡儿积 $A\times B$ 是什么？

解 笛卡儿积 $A\times B$ 由形为 (a, b) 的所有有序偶组成，其中 a 是个学生，而 b 是该校开的一门课。集合 $A\times B$ 可以用来表示该校学生选课的所有可能情况。■

例 14 什么是 $A=\{1, 2\}$ 和 $B=\{a, b, c\}$ 的笛卡儿积？

解 笛卡儿积 $A\times B$ 是

$$A\times B=\{(1,a),(1,b),(1,c),(2,a),(2,b),(2,c)\}$$ ■

笛卡儿积 $A\times B$ 的子集 R 称为从集合 A 到集合 B 的关系。R 的元素是有序对，其中第一个元素属于 A，第二个因素属于 B。例如，$R=\{(a, 0), (a, 1), (a, 3), (b, 1), (b, 2), (c, 0), (c, 3)\}$ 是从集合 $\{a, b, c\}$ 到集合 $\{0, 1, 2, 3\}$ 的关系。我们将在第 7 章研究关系。

除非 $A=\varnothing$ 或 $B=\varnothing$（这样 $A\times B=\varnothing$）或除非 $A=B$，否则 $A\times B$ 和 $B\times A$ 不相等（参看本节末尾练习 26）。这在下例中说明。

例 15 证明笛卡儿积 $B\times A$ 和笛卡儿积 $A\times B$ 不相等，其中 A 和 B 为例 14 中的集合。

解 笛卡儿积 $B\times A$ 是

$$B\times A=\{(a,1),(a,2),(b,1),(b,2),(c,1),(c,2)\}$$

这不等于例 14 中得到的 $A\times B$。■

对于两个以上的集合也可以定义笛卡儿积。

定义 10 集合 $A_1, A_2, \cdots, A_n$ 的笛卡儿积用 $A_1\times A_2\times\cdots\times A_n$ 表示，这是有序 n 元组 $(a_1, a_2, \cdots, a_n)$ 的集合，其中对于 $i=1, 2, \cdots, n$，$a_i\in A_i$。换言之，

$$A_1\times A_2\times\cdots\times A_n=\{(a_1,a_2,\cdots,a_n)\mid a_i\in A_i, i=1,2,\cdots,n\}$$

例 16 什么是笛卡儿积 $A\times B\times C$，其中 $A=\{0, 1\}$，$B=\{1, 2\}$，$C=\{0, 1, 2\}$？

解 这一笛卡儿积由所有有序三元组 (a, b, c) 组成，其中 $a\in A$，$b\in B$，$c\in C$。因此，

$$A\times B\times C=\{(0,1,0),(0,1,1),(0,1,2),(0,2,0),(0,2,1),(0,2,2),(1,1,0),(1,1,1),(1,1,2),(1,2,0),(1,2,1),(1,2,2)\}$$ ■

1.6.4 使用带量词的集合符号

有时我们在符号中显式说明语句的论域。特别是，$\forall x\in SP(x)$ 表示 $P(x)$ 的全称量化，其中论域是集合 S。类似地，$\exists x\in SP(x)$ 表示 $P(x)$ 的存在量化，论域是 S。

例 17 语句 $\forall x\in\mathbf{R}(x^2\geqslant 0)$ 和 $\exists x\in\mathbf{Z}(x^2=1)$ 的含义是什么？

解 语句 $\forall x\in\mathbf{R}(x^2\geqslant 0)$ 声称对任意实数 x，$x^2\geqslant 0$。这个语句可以表达为“任意实数的平方是非负的”。这是真语句。

语句 $\exists x\in\mathbf{Z}(x^2=1)$ 声称存在一个整数 x 使得 $x^2=1$。这个语句可以表达为“有某个整数，其平方是 1”。这个语句也是真语句，因为 $x=1$ 就是这样一个整数（-1 也是）。■

练习

1. 列出下述集合的成员。

   a) $\{x\mid x$ 是使得 $x^2=1$ 的实数$\}$

   b) $\{x\mid x$ 是小于 12 的正整数$\}$

   c) $\{x\mid x$ 是某个整数的平方且 $x<100\}$

   d) $\{x\mid x$ 是整数且 $x^2=2\}$

2. 用集合构造符号，给出下列每个集合的描述。

a) $\{0, 3, 6, 9, 12\}$　　b) $\{-3, -2, -1, 0, 1, 2, 3\}$

c) $\{m, n, o, p\}$

3. 判断下面每对集合是否相等。

a) $\{1, 3, 3, 3, 5, 5, 5, 5, 5\}$，$\{5, 3, 1\}$

b) $\{\{1\}\}$，$\{1, \{1\}\}$

c) $\varnothing$，$\{\varnothing\}$

4. 设 $A=\{2, 4, 6\}$，$B=\{2, 6\}$，$C=\{4, 6\}$，$D=\{4, 6, 8\}$。判断这些集合中哪个是另外一个的子集。

5. 对下面的每个集合，判断 2 是否为它的元素。

a) $\{x\in\mathbf{R} \mid x$ 是大于 1 的整数$\}$　　b) $\{x\in\mathbf{R} \mid x$ 是某整数的平方$\}$

c) $\{2, \{2\}\}$　　d) $\{\{2\}, \{\{2\}\}\}$

e) $\{\{2\}, \{2, \{2\}\}\}$　　f) $\{\{\{2\}\}\}$

6. 对练习 5 中的每个集合，判断$\{2\}$是否它的一个元素。

7. 判断下列语句是真还是假。

a) $0\in\varnothing$　　b) $\varnothing\in\{0\}$　　c) $\{0\}\subset\varnothing$

d) $\varnothing\subset\{0\}$　　e) $\{0\}\in\{0\}$　　f) $\{0\}\subset\{0\}$

g) $\{\varnothing\}\subseteq\{\varnothing\}$

8. 判断下列语句是真还是假。

a) $\varnothing\in\{\varnothing\}$　　b) $\varnothing\in\{\varnothing, \{\varnothing\}\}$　　c) $\{\varnothing\}\in\{\varnothing\}$

d) $\{\varnothing\}\in\{\{\varnothing\}\}$　　e) $\{\varnothing\}\subset\{\varnothing, \{\varnothing\}\}$　　f) $\{\{\varnothing\}\}\subset\{\varnothing, \{\varnothing\}\}$

g) $\{\{\varnothing\}\}\subset\{\{\varnothing\}, \{\varnothing\}\}$

9. 判断下列语句是真还是假。

a) $x\in\{x\}$　　b) $\{x\}\subseteq\{x\}$　　c) $\{x\}\in\{x\}$

d) $\{x\}\in\{\{x\}\}$　　e) $\varnothing\subseteq\{x\}$　　f) $\varnothing\in\{x\}$

10. 用文氏图说明集合关系 $A\subseteq B$ 和 $B\subseteq C$。

11. 假定 A、B 和 C 为集合，且 $A\subseteq B$，$B\subseteq C$。证明 $A\subseteq C$。

12. 找出两个集合 A 和 B，使得 $A\in B$ 且 $A\subseteq B$。

13. 下列各集合的基数是什么？

a) $\{a\}$　　b) $\{\{a\}\}$　　c) $\{a, \{a\}\}$

d) $\{a, \{a\}, \{a, \{a\}\}\}$

14. 下列各集合的基数是什么？

a) $\varnothing$　　b) $\{\varnothing\}$　　c) $\{\varnothing, \{\varnothing\}\}$

d) $\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$

15. 找出下列各集合的幂集合。

a) $\{a\}$　　b) $\{a, b\}$　　c) $\{\varnothing, \{\varnothing\}\}$

16. 如果 A 和 B 是两个集合，且有相同的幂集合，能否肯定 $A=B$？

17. 下列各集合各有多少个元素？

a) $P(\{a, b, \{a, b\}\})$　　b) $P(\{\varnothing, a, \{a\}, \{\{a\}\}\})$　　c) $P(P(\varnothing))$

18. 判断下列各集合是否为某集合的幂集合。

a) $\varnothing$　　b) $\{\varnothing, \{a\}\}$　　c) $\{\varnothing, \{a\}, \{\varnothing, a\}\}$

d) $\{\varnothing, \{a\}, \{b\}, \{a, b\}\}$

19. 令 $A=\{a, b, c, d\}$，$B=\{y, z\}$。求

a) $A\times B$　　b) $B\times A$

20. 令 A 为某大学的数学系开设课程的集合，B 为该大学所有数学教授的集合，笛卡儿集 $A\times B$ 是什么？

21. 什么是笛卡儿集 $A\times B\times C$，其中 A 是所有航线的集合，B 和 C 都是所有美国城市的集合。

22. 假定 $A\times B=\varnothing$，其中 A 和 B 为集合，由此你得出什么结论？

23. 令 A 为集合，求证 $\varnothing\times A=A\times\varnothing=\varnothing$。

24. 令 $A=\{a, b, c\}$，$B=\{x, y\}$，$C=\{0, 1\}$。求

a) $A\times B\times C$ b) $C\times B\times A$

c) $C\times A\times B$ d) $B\times B\times B$

25. 如果 A 有 m 个元素，B 有 n 个元素，$A\times B$ 有多少个元素？

26. 求证除非 $A=B$，否则 $A\times B\neq B\times A$，其中 A 和 B 均为非空集合。

27. 将下列量化表达式翻译成汉语句子并确定其真值。

a) $\forall x\in \mathbf{R}(x^2\neq -2)$ b) $\exists x\in \mathbf{Z}(x^2=2)$

c) $\forall x\in \mathbf{Z}(x^2>0)$ d) $\exists x\in \mathbf{R}(x^2=x)$

28. 将下列量化表达式翻译成汉语句子并确定其真值。

a) $\exists x\in \mathbf{R}(x^3=-1)$ b) $\exists x\in \mathbf{Z}(x+1>x)$

c) $\forall x\in \mathbf{Z}(x-1\in \mathbf{Z})$ d) $\forall x\in \mathbf{Z}(x^2\in \mathbf{Z})$

*29. 证明序偶 (a, b) 可以用集合术语定义为 $\{\{a\}, \{a, b\}\}$。[提示：首先证明 $\{\{a\}, \{a, b\}\}=\{\{c\}, \{c, d\}\}$ 当且仅当 $a=c$ 且 $b=d$。]

*30. 本练习介绍罗素悖论。令集合 S 为包含所有不以自身为元素的那些集合，即 $S=\{x \mid x\notin x\}$。

a) 证明从 S 是它自己的一个元素的假设能推出矛盾。

b) 证明从 S 不是它自己的一个元素的假设能推出矛盾。

从 a) 和 b) 知 S 不可能是其定义中给出的那个集合。不过这一悖论是可以避免的，只要对可以作为集合元素的类型加以限制即可。

*31. 给出可以列出有限集合所有子集的步骤。

1.7 集合运算

1.7.1 引言

两个集合可以以许多不同的方式结合在一起。例如，从学校主修数学课的集合和主修计算机科学课的集合入手，可以构成主修数学或计算机科学的学生集合，既主修数学又主修计算机科学的学生集合，不主修数学的学生集合，等等。

定义 1 令 A 和 B 为集合。A 和 B 的并集用 $A\cup B$ 表示，这是 A 或 B 中或同时在 A 和 B 中的元素组成的集合。

元素 x 属于 A 和 B 的并集当且仅当 x 属于 A 或 x 属于 B。这说明

$$A\cup B=\{x \mid x\in A \vee x\in B\}$$

图 1-3 中的文氏图表示集合 A 和 B 的并集。代表集合 A 的圆圈内和代表 B 的圆圈内的阴影区域表示 A 和 B 的并集。

我们将给出集合并集的例子。

例 1 集合 $\{1, 3, 5\}$ 和集合 $\{1, 2, 3\}$ 的并集是集合 $\{1, 2, 3, 5\}$，即

$$\{1,3,5\}\cup\{1,2,3\}=\{1,2,3,5\}$$ ■

例 2 学校主修计算机科学的学生集合与主修数学的学生集合的并集，是或主修数学、或主修计算机科学或同时主修这两门课的学生的集合。 ■

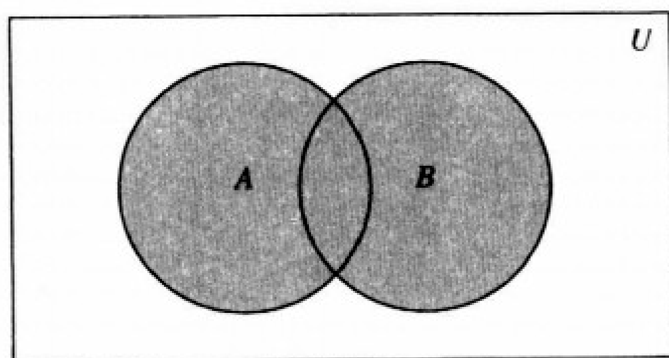
定义 2 令 A 和 B 的集合。A 和 B 的交集用 $A\cap B$ 表示，这是既在 A 中又在 B 中的那些元素的集合。

元素 x 属于集合 A 和 B 的交集当且仅当 x 属于 A 而且 x 属于 B。这说明

$$A\cap B=\{x \mid x\in A \wedge x\in B\}$$

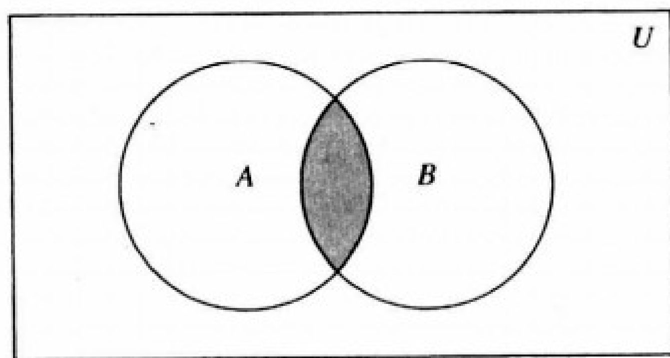
图 1-4 中的文氏图表示集合 A 和 B 的交集。同时在代表 A 和 B 的两个园之内的阴影区域表示 A 和 B 的交集。

我们给出交集的几个例子。



$A \cup B$为阴影区

图 1-3 表示 A 和 B 并集的文氏图



$A \cap B$为阴影区

图 1-4 表示 A 和 B 交集的文氏图

例 3 集合$\{1, 3, 5\}$和$\{1, 2, 3\}$的交集是$\{1, 3\}$，即$\{1, 3, 5\} \cap \{1, 2, 3\} = \{1, 3\}$。■

例 4 学校所有主修计算机科学的学生集合与所有主修数学的学生集合，是所有既主修计算机科学又主修数学的学生的集合。■

定义 3 如果两个集合的交集为空集，就说它们不相交。

例 5 令 $A=\{1, 3, 5, 7, 9\}$而 $B=\{2, 4, 6, 8, 10\}$。由于 $A \cap B=\varnothing$，A 和 B 不相交。■

我们往往有兴趣求出集合的并集的基数。要计算两个有限集合 A 和 B 的并集的元素个数，须注意 $|A|+|B|$ 把只属于 A 或只属于 B 的元素数了一次，而既属于 A 又属于 B 的元素数了两次。因此，如果从 $|A|+|B|$ 中减去既属于 A 又属于 B 的元素的个数，则 $A \cap B$ 中的元素也就只数一次。于是

$$|A \cup B| = |A| + |B| - |A \cap B|$$

把这一结果推广到任意多个集合的并集，就得到所谓的*包含排斥原理*(或简称容斥原理)。容斥原理是用于枚举的一项重要技术。我们将在第 4 章和第 6 章详细讨论这一原理和其他的计数技术。

还有其他组合集合的重要方式。

定义 4 令 A 和 B 为集合。A 和 B 的差集用 $A-B$ 表示，这是包含只属于 A 而不属于 B 的所有元素的集合。A 和 B 的差集也称为 B 对于 A 的补集。

元素 x 属于 A 和 B 的差集当且仅当 $x \in A$ 且 $x \notin B$，这说明

$$A - B = \{x \mid x \in A \land x \notin B\}$$

图 1-5 中的文氏图表示集合 A 和 B 的差集。既在表示集合 A 的圆圈内部也在表示集合 B 的圆圈外部的阴影区域表示 $A-B$。

让我们举几个差集的例子。

例 6 集合$\{1, 3, 5\}$和$\{1, 2, 3\}$的差集是$\{5\}$；即$\{1, 3, 5\}-\{1, 2, 3\}=\{5\}$。这不同于$\{1, 2, 3\}$和$\{1, 3, 5\}$的差集$\{2\}$。■

例 7 学校主修计算机科学的学生集合和主修数学的学生集合的差集是学校主修计算机科学但不主修数学的学生集合。■

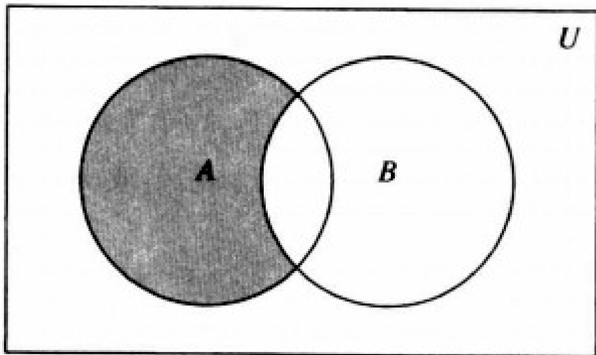
一旦指定了全集 U，就可以定义集合的补集。

定义 5 令 U 为全集。集合 A 的补集用 $\overline{A}$ 表示，这是 A 对于 U 的补集。集合 A 的补集是$U-A$。

元素 x 属于 $\overline{A}$ 当且仅当 $x \notin A$。这说明

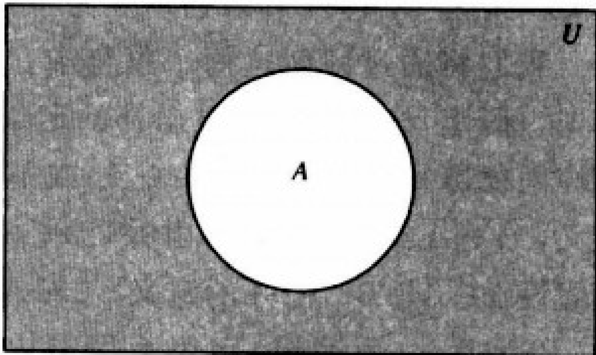
$$\overline{A} = \{x \mid x \notin A\}$$

图 1-6 中代表集合 A 的圆圈外面的阴影区域表示 $\overline{A}$。



A − B 为阴影区

图 1-5 A 和 B 的差集的文氏图



$\overline{A}$ 为阴影区

图 1-6 集合 A 的补集的文氏图

我们举几个补集的例子。

例 8 令 $A=\{a, e, i, o, u\}$（全集为英语字母集）。那么

$$\overline{A} = \{b,c,d,f,g,h,j,k,l,m,n,p,q,r,s,t,v,w,x,y,z\}$$ ■

例 9 令 A 为大于 10 的正整数的集合（全集为正整数集合）。那么

$$\overline{A} = \{1,2,3,4,5,6,7,8,9,10\}$$ ■

1.7.2 集合恒等式

表 1-21 列出了最重要的集合恒等式。本节将用三种不同的方法证明其中的几个。介绍这些方法是想说明，对一个问题的解往往有不同的途径。表中未证明的恒等式留给读者作练习。读者应该注意这些集合恒等式和 1.2 节讨论的逻辑等价的相似之处。事实上，这里给出的集合恒等式可以直接由对应的逻辑等价证明。不仅如此，而且这两者都是布尔代数（在第 10 章讨论）中成立的恒等式的特殊情况。

表 1-21 集合恒等式

等 式	名 称
$A\cup\varnothing=A$ $A\cap U=A$	恒等律
$A\cup U=U$ $A\cap\varnothing=\varnothing$	支配律
$A\cup A=A$ $A\cap A=A$	幂等律
$\overline{(\overline{A})}=A$	补集律
$A\cup B=B\cup A$ $A\cap B=B\cap A$	交换律
$A\cup(B\cup C)=(A\cup B)\cup C$ $A\cap(B\cap C)=(A\cap B)\cap C$	结合律
$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$ $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$	分配律
$\overline{A\cup B}=\overline{A}\cap\overline{B}$ $\overline{A\cap B}=\overline{A}\cup\overline{B}$	德摩根定律
$A\cup(A\cap B)=A$ $A\cap(A\cup B)=A$	吸收律
$A\cup\overline{A}=U$ $A\cap\overline{A}=\varnothing$	补律

Extra Examples 证明集合相等的一种方法是证明两者中任何一个都是另一个的子集。我们将以确立德摩根第二定律为例说明这一方法。

例 10 用证明$\overline{A\cap B}$和$\bar{A}\cup\bar{B}$互为子集的方法证明$\overline{A\cap B}=\bar{A}\cup\bar{B}$。

解 我们将通过证明这两个集合中，一个是另一个的子集来证明这两个集合是相等的。

首先，假定$x\in\overline{A\cap B}$。根据补的定义，$x\notin A\cap B$。由交的定义，$\neg((x\in A)\wedge(x\in B))$为真。(由逻辑)用德摩根定律可得$\neg(x\in A)$或$\neg(x\in B)$。因此，根据逆的定义，有$x\notin A$或$x\notin B$。再由补的定义，$x\in\bar{A}$或$x\in\bar{B}$。再由并，$x\in\bar{A}\cup\bar{B}$。因此，得证$\overline{A\cap B}\leqslant\bar{A}\cup\bar{B}$。

现假设$x\in\bar{A}\cup\bar{B}$。由并的定义，$x\in\bar{A}$或$x\in\bar{B}$。用补的定义，$x\in A$或$x\notin B$。所以，$\neg(x\in A)\vee\neg(x\in B)$为真。(由逻辑)用德摩根定律，得$\neg((x\in A)\wedge(x\in B))$为真。由交的定义，$\neg(x\in A\cap B)$成立。由补的定义，有$x\in\overline{A\cap B}$。这表明$\bar{A}\cup\bar{B}\subseteq\overline{A\cap B}$。由于已证明这两个集合中，一个是另一个的子集，所以，这两个集合相等，即得证恒等。 ■

另一个更简洁的方法是使用集合构造符和例 10 中的推理。如例 11 所示。

例 11 用集合构造符和逻辑等价证明$\overline{A\cap B}=\bar{A}\cup\bar{B}$。

解 下列一串等式提供了这一恒等关系的证明：

$$\begin{aligned}\overline{A\cap B}&=\{x\mid x\notin A\cap B\}\\&=\{x\mid\neg(x\in(A\cap B))\}\\&=\{x\mid\neg(x\in A\wedge x\in B)\}\\&=\{x\mid x\notin A\vee x\notin B\}\\&=\{x\mid x\in\bar{A}\vee x\in\bar{B}\}\\&=\{x\mid x\in\bar{A}\cup\bar{B}\}\\&=\bar{A}\cup\bar{B}\end{aligned}$$

注意在这一串等式中的第四个等号处使用了逻辑等价的德摩根第二定律。 ■

要证明涉及两个以上集合的集合恒等，可以证明恒等式的每一边是另一边的子集，这我们可用例 12 来说明。

例 12 对任意集合A，B和C，证明$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$。

解 我们将通过说明等式的每一边是另一边的子集来证明这个恒等式。

假定$x\in A\cap(B\cup C)$，那么$x\in A$且$x\in B\cup C$。根据集合并的定义，可得$x\in A$且$x\in B$或$x\in C$。因此，$x\in A$且$x\in B$或$x\in A$且$x\in C$。根据交的定义，可知$x\in A\cap B$或$x\in A\cap C$。使用并的定义，可得出$x\in(A\cap B)\cup(A\cap C)$。由此$A\cap(B\cup C)\subseteq(A\cap B)\cup(A\cap C)$。

现在假定$x\in(A\cap B)\cup(A\cap C)$。那么，由并的定义，$x\in A\cap B$或$x\in A\cap C$。由交的定义，可得$x\in A$且$x\in B$或$x\in A$且$x\in C$。由此可知$x\in A$，$x\in B$或$x\in C$。因此，由并的定义可知$x\in A$且$x\in B\cup C$。而且，由交的定义，可得$x\in A\cap(B\cup C)$。最终得出$(A\cap B)\cup(A\cap C)\subseteq A\cap(B\cup C)$。这便完成了恒等式的证明。 ■

还可以用成员表来证明集合恒等式。我们考虑一个元素可能属于的集合的每一种组合，并证明在同样的集合组合中的元素属于恒等式两边的集合。用 1 表示元素属于一个集合，用 0 表示元素不属于一个集合(读者应注意成员表和真值表的相似之处)。

例 13 用成员表证明$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$。

解 表 1-22 给出了这些集合组合的成员表。这表格有 8 行。由于对应于$A\cap(B\cup C)$和$(A\cap B)\cup(A\cap C)$的两列相同，恒等式有效。 ■

用我们已经证明的这些集合恒等式可以证明其他的集合恒等式。考虑下面的例子。

表 1-22 分配性质的成员表

A	B	C	$B\cup C$	$A\cap(B\cup C)$	$A\cap B$	$A\cap C$	$(A\cap B)\cup(A\cap C)$
1	1	1	1	1	1	1	1
1	1	0	1	1	1	0	1
1	0	1	1	1	0	1	1
1	0	0	0	0	0	0	0
0	1	1	1	0	0	0	0
0	1	0	1	0	0	0	0
0	0	1	1	0	0	0	0
0	0	0	0	0	0	0	0

例 14 令 A，B，C 为集合。证明

$$\overline{A\cup(B\cap C)}=(\overline{C}\cup\overline{B})\cap\overline{A}$$

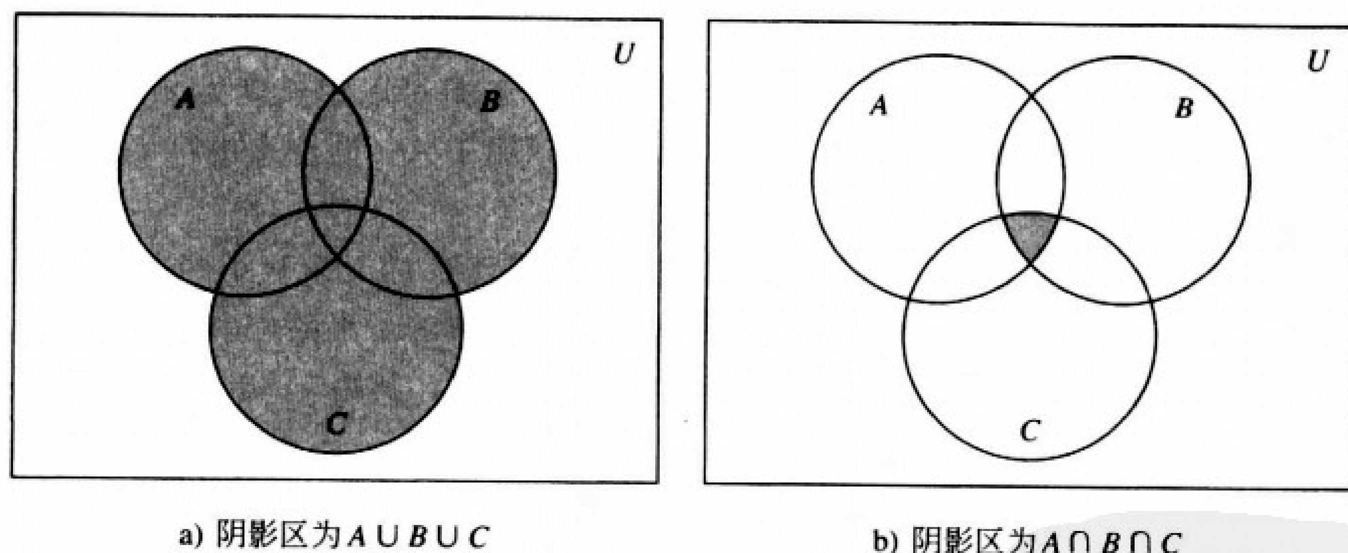
解 我们有

$$\begin{aligned}\overline{A\cup(B\cap C)}&=\overline{A}\cap(\overline{B\cap C}) && \text{德摩根第一律}\\&=\overline{A}\cap(\overline{B}\cup\overline{C}) && \text{德摩根第二律}\\&=(\overline{B}\cup\overline{C})\cap\overline{A} && \text{交集的交换律}\\&=(\overline{C}\cup\overline{B})\cap\overline{A} && \text{并集的交换律}\end{aligned}$$

■

1.7.3 扩展的并集和交集

由于集合的并集和交集满足结合律，所以只要 A、B、C 为集合，$A\cup B\cup C$ 和 $A\cap B\cap C$ 均有定义。注意 $A\cup B\cup C$ 包含那些至少属于 A、B、C 中一个集合的元素，而 $A\cap B\cap C$ 包含那些属于 A、B、C 全部三个集合的元素。三个集合 A、B、C 的这两种组合如图 1-7 所示。



a) 阴影区为 $A\cup B\cup C$　　b) 阴影区为 $A\cap B\cap C$

图 1-7 集合 A、B、C 的并集和交集

例 15 令 $A=\{0, 2, 4, 6, 8\}$，$B=\{0, 1, 2, 3, 4\}$，$C=\{0, 3, 6, 9\}$。$A\cup B\cup C$ 和 $A\cap B\cap C$ 是什么集合？

解 $A\cup B\cup C$ 包括那些至少属于 A，B，C 之一的元素。所以

$$A\cup B\cup C=\{0,1,2,3,4,6,8,9\}$$

集合 $A\cap B\cap C$ 包括那些属于全部三个集合的元素。因此

$$A\cap B\cap C=\{0\}$$

■

可以考虑任意多个集合的并集和交集。我们采用下面的定义。

定义 6 一组集合的并集是包含那些至少是这组集合中一个集合成员的元素的集合。

用记号

$$A_1 \cup A_2 \cup \cdots \cup A_n = \bigcup_{i=1}^{n} A_i$$

表示集合 A_1，A_2，…，A_n 的并集。

定义 7 一组集合的交集是包含那些属于这组集合中所有成员集合的元素的集合。

用记号

$$A_1 \cap A_2 \cap \cdots \cap A_n = \bigcap_{i=1}^{n} A_i$$

表示集合 A_1，A_2，…，A_n 的交集。我们用例 16 说明扩展的并集和交集。

例 16 令 $A_i=\{i, i+1, i+2, \cdots\}$。那么

$$\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n} \{i,i+1,i+2,\cdots\} = \{1,2,3,\cdots\}$$

而

$$\bigcap_{i=1}^{n} A_i = \bigcap_{i=1}^{n} \{i,i+1,i+2,\cdots\} = \{n,n+1,n+2,\cdots\}$$ ■

1.7.4 计算机表示集合的方式

计算机表示集合的方式各种各样。一种办法是把集合的元素无序地存储起来。可是如果这样做，在做集合的并集、交集或差集等运算时会浪费时间，因为这些运算将需要大量的元素检索。我们将要介绍一种利用全集元素的一个任意排序存放元素以表示集合的方法。集合的这种表示法使我们很容易计算集合的组合。

假定全集 U 是有限的(而且大小合适，使 U 的元素个数不超过计算机能使用的内存量)。首先为 U 的元素任意规定一个顺序，例如 a_1，a_2，…，a_n。于是可以用长度为 n 的位串表示 U 的子集 A：如果 a_i 属于 A，则位串中第 i 位是 1；如果 a_i 不属于 A，则位串中第 i 位是 0。下面的例子阐明了这一方法。

例 17 令 $U=\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$，而且 U 的元素从小到大排序，即 $a_i=i$。表示 U 中所有奇数的子集、所有偶数的子集和不超过 5 的整数的子集的位串是什么？

解 表示 U 中所有奇数的子集即$\{1, 3, 5, 7, 9\}$的位串，其第 1、3、5、7、9 位为 1，其他位为 0，即

10 1010 1010

(我们已把长度为 10 的位串分成长度为 5 的两段以便阅读，因为长位串不易读)。类似地，U 中所有偶数的子集即$\{2, 4, 5, 8, 10\}$，由位串

01 0101 0101

表示。U 中不超过 5 的所有整数的集合即$\{1, 2, 3, 4, 5\}$，由位串

11 1110 0000

表示。 ■

用位串表示集合便于计算集合的补集、并集、交集和差集。要从表示集合的位串计算它的补集的位串，只须简单地把每个 1 改为 0，每个 0 改为 1，因为 $x \in A$ 当且仅当 $x \notin \overline{A}$。注意，当把每个字位看成是真值时，上述运算对应于取每个字位的非，因为 1 表示真，0 表示假。

例 18 我们已经知道集合$\{1, 3, 5, 7, 9\}$的位串(全集为$\{1, 2, 3, 4, 5, 6, 7, 8, 9,$

10})是

$$10\ 1010\ 1010$$

它的补集的位串是什么？

解 用0取代1，用1取代0，即可得到此集合的补集的位串

$$01\ 0101\ 0101$$

这对应着集合{2，4，6，8，10}。 ■

要得到两个集合的并集和交集的位串，我们可以对表示这两个集合的位串按位做布尔运算。只要两个位串的第 i 字位有一个是1，则并集的位串的第 i 位是1，当两个字位都是0时为0。因此，并集的位串是两个集合位串的按位或(bitwise *OR*)。当两个位串的第 i 字位均为1时，交集的位串第 i 位为1，否则为0。因此交集的位串是两个集合位串的按位与(bitwise *AND*)。

例 19 集合{1，2，3，4，5}和{1，3，5，7，9}的位串分别是11 1110 0000和10 1010 1010。用位串找出它们的并集和交集。

解 这两个集合的并集的位串是

$$11\ 1110\ 0000 \vee 10\ 1010\ 1010 = 11\ 1110\ 1010$$

它表示的集合是{1，2，3，4，5，7，9}。这两个集合的交集的位串是

$$11\ 1110\ 0000 \wedge 10\ 1010\ 1010 = 10\ 1010\ 0000$$

它表示的集合是{1，3，5}。 ■

练习

1. 令 A 为住在离学校一英里以内的所有学生的集合，B 是走路上学的所有学生的集合。描述一下下列各集合中的学生：
   a) $A\cap B$　　b) $A\cup B$　　c) $A-B$　　d) $B-A$
2. 假定 A 是学校二年级的学生集合，B 是学校上离散数学课的学生集合。用 A 和 B 来表示下列各个集合。
   a)学校二年级上离散数学课的学生集合。
   b)学校二年级不上离散数学课的学生集合。
   c)学校二年级的或上离散数学课的学生集合。
   d)学校既不在二年级的也不上离散数学课的学生集合。
3. 令 $A=\{1, 2, 3, 4, 5\}$，$B=\{0, 3, 6\}$。求
   a) $A\cup B$　　b) $A\cap B$　　c) $A-B$　　d) $B-A$
4. 令 $A=\{a, b, c, d, e\}$，$B=\{a, b, c, d, e, f, g, h\}$。求
   a) $A\cup B$　　b) $A\cap B$　　c) $A-B$　　d) $B-A$
5. 令 A 为集合。证明 $\overline{\overline{A}}=A$。
6. 令 A 为集合。证明
   a) $A\cup\varnothing=A$　　b) $A\cap\varnothing=\varnothing$　　c) $A\cup A=A$　　d) $A\cap A=A$
   e) $A-\varnothing=A$　　f) $A\cup U=U$　　g) $A\cap U=A$　　h) $\varnothing-A=\varnothing$
7. 令 A 和 B 为集合。证明
   a) $A\cup B=B\cup A$　　b) $A\cap B=B\cap A$
8. 令 A 和 B 为集合，证明 $A\cup(A\cap B)=A$。
9. 令 A 和 B 为集合，证明 $A\cap(A\cup B)=A$。
10. 若 $A-B=\{1, 5, 7, 8\}$，$B-A=\{2, 10\}$，而 $A\cap B=\{3, 6, 9\}$，求集合 A 和 B。
11. 求证若 A 和 B 为集合，则 $\overline{A\cup B}=\overline{A}\cap\overline{B}$。
    a)用证明等式两边互为子集的办法。　　b)用成员表。
12. 令 A，B 为集合。证明

a) $(A\cap B)\subseteq A$　b) $A\subseteq(A\cup B)$　c) $A-B\subseteq A$　d) $A\cap(B-A)=\varnothing$
e) $A\cup(B-A)=A\cup B$

13. 设 A，B，C 为集合。证明 $\overline{A\cap B\cap C}=\overline{A}\cup\overline{B}\cup\overline{C}$。
a)用证明等式两边互为子集的办法。　b)用成员表。

14. 令 A，B，C 为集合。证明
a) $(A\cup B)\subseteq(A\cup B\cup C)$　b) $(A\cap B\cap C)\subseteq(A\cap B)$
c) $(A-B)-C\subseteq A-C$　d) $(A-C)\cap(C-B)=\varnothing$
e) $(B-A)\cup(C-A)=(B\cup C)-A$

15. 如果 A 和 B 为集合，证明 $A-B=A\cap\overline{B}$。

16. 如果 A 和 B 为集合，证明 $(A\cap B)\cup(A\cap\overline{B})=A$。

17. 令 A，B，C 为集合。证明
a) $A\cup(B\cup C)=(A\cup B)\cup C$　b) $A\cap(B\cap C)=(A\cap B)\cap C$
c) $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$

18. 令 A，B，C 为集合。证明 $(A-B)-C=(A-C)-(B-C)$。

19. 令 $A=\{0, 2, 4, 6, 8, 10\}$，$B=\{0, 1, 2, 3, 4, 5, 6\}$，$C=\{4, 5, 6, 7, 8, 9, 10\}$。求
a) $A\cap B\cap C$　b) $A\cup B\cup C$　c) $(A\cup B)\cap C$　d) $(A\cap B)\cup C$

20. 画出集合 A，B，C 的下列组合的文氏图。
a) $A\cap(B\cup C)$　b) $\overline{A}\cap\overline{B}\cap\overline{C}$　c) $(A-B)\cup(A-C)\cup(B-C)$

21. 如果有关集合 A 与 B 的下列性质为真，你能就 A 和 B 说些什么？
a) $A\cup B=A$　b) $A\cap B=A$　c) $A-B=A$
d) $A\cap B=B\cap A$　e) $A-B=B-A$

22. 如果集合 A，B，C 满足下述条件，能断定 $A=B$ 吗？
a) $A\cup C=B\cup C$　b) $A\cap C=B\cap C$

23. 令 A 和 B 为全集 U 的子集。证明 $A\subseteq B$ 当且仅当 $\overline{B}\subseteq\overline{A}$。

集合 A 和 B 的对称差，用 $A\oplus B$ 表示，是属于 A 或属于 B 但不同时属于 A 与 B 的元素组成的集合。

24. 求 $\{1, 3, 5\}$ 和 $\{1, 2, 3\}$ 的对称差。

25. 求某校主修计算机科学的学生集合与主修数学的学生集合的对称差。

26. 画出集合 A 与 B 的对称差的文氏图。

27. 证明 $A\oplus B=(A\cup B)-(A\cap B)$。

28. 证明 $A\oplus B=(A-B)\cup(B-A)$。

29. 证明：若 A 是全集 U 的子集，则
a) $A\oplus A=\varnothing$　b) $A\oplus\varnothing=A$　c) $A\oplus U=\overline{A}$　d) $A\oplus\overline{A}=U$

30. 如果 A 和 B 为集合，证明
a) $A\oplus B=B\oplus A$　b) $(A\oplus B)\oplus B=A$

31. 如果 $A\oplus B=A$，你能就集合 A 和 B 说些什么？

*32. 判断对称差是否满足结合律；即若 A，B，C 为集合，

$$A\oplus(B\oplus C)=(A\oplus B)\oplus C$$

是否成立？

*33. 假定 A，B，C 为集合，使得 $A\oplus C=B\oplus C$。是否必定有 $A=B$？

34. 若 A，B，C，D 为集合，$(A\oplus B)\oplus(C\oplus D)=(A\oplus C)\oplus(B\oplus D)$ 是否成立？

35. 若 A，B，C，D 为集合，$(A\oplus B)\oplus(C\oplus D)=(A\oplus D)\oplus(B\oplus C)$ 是否成立？

*36. 证明：若 A，B，C 为集合，则

$$|A\cup B\cup C|=|A|+|B|+|C|-|A\cap B|-|A\cap C|-|B\cap C|+|A\cap B\cap C|$$

(这是第6章将要学习的包含排斥原理的一个特例。)

*37. 令 $A_i=\{1, 2, 3, \cdots, i\}$，$i=1, 2, 3, \cdots$。求

a) $\bigcup_{i=1}^{n} A_i$　　b) $\bigcap_{i=1}^{n} A_i$

*38. 令 $A_i=\{\cdots, -2, -1, 0, 1, \cdots, i\}$。求

a) $\bigcup_{i=1}^{n} A_i$　　b) $\bigcap_{i=1}^{n} A_i$

39. 令 A_i 为所有长度不超过 i 的非空位串(即长度至少为 1)的集合。求

a) $\bigcup_{i=1}^{n} A_i$　　b) $\bigcap_{i=1}^{n} A_i$

40. 假定全集 $U=\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$。用位串表示下列各集合，假定若 i 属于该集合则其位串第 i 位为 1，否则第 i 位为 0。

a) $\{3, 4, 5\}$　　b) $\{1, 3, 6, 10\}$　　c) $\{2, 3, 4, 7, 8, 9\}$

41. 使用上题中的同一个全集，求下列位串各自代表的集合。

a) 11 1100 1111　　b) 01 0111 1000　　c) 10 0000 0001

42. 下列位串各代表有限全集的什么子集？

a) 所有字位全为 0 的串。　　b) 所有字位全为 1 的串。

43. 对应于两个集合之差的位串是什么？

44. 对应于两个集合的对称差的位串是什么？

45. 令 $A=\{a, b, c, d, e\}$，$B=\{b, c, d, g, p, t, v\}$，$C=\{c, e, i, o, u, x, y, z\}$，$D=\{d, e, h, i, n, o, t, u, x, y\}$。说明怎样用位串的按位运算求下列集合：

a) $A\cup B$　　b) $A\cap B$　　c) $(A\cup D)\cap(B\cup C)$　　d) $A\cup B\cup C\cup D$

46. 怎样用位串求出同一全集 U 的 n 个子集的并集和交集？

47. 集合 A 的后继是 $A\cup\{A\}$。求下列集合的后继。

a) $\{1, 2, 3\}$　　b) $\varnothing$　　c) $\{\varnothing\}$　　d) $\{\varnothing, \{\varnothing\}\}$

48. 含 n 个元素的集合的后继有几个元素？

有时一个元素在一个无序集中出现的次数也有意义。当同一个元素作为成员可以出现不止一次时，这个无序元素集就是多重集。符号 $\{m_1\cdot a_1, m_2\cdot a_2, \cdots, m_r\cdot a_r\}$ 表示的是 a_1 出现 m_1 次，a_2 出现 m_2 次等的多重集。整数 m_i 称为 a_i 的重数，$i=1, 2, \cdots, r$。

令 P 和 Q 为多重集，多重集 P 和 Q 的并集是个多重集，其中每个元素的重数是该元素在 P 和 Q 中的重数的最大值。P 和 Q 的交集是个多重集，其中每个元素的重数是该元素在 P 和 Q 中重数的最小值。P 和 Q 的差集是个多重集，其中每个元素的重数是该元素在 P 中的重数减去它在 Q 中的重数，当然这不能是个负数，若是负数就以 0 作重数。P 和 Q 的和集是个多重集，其中每个元素的重数是该元素在 P 和 Q 中的重数的和，P 和 Q 的并集，交集和差集分别用 $P\cup Q$，$P\cap Q$ 和 $P-Q$ 表示(不要混淆这些运算和集合上的类似运算)。P 和 Q 的和集用 $P+Q$ 表示。

49. 令 A 和 B 分别为多重集 $\{3\cdot a, 2\cdot b, 1\cdot c\}$ 和 $\{2\cdot a, 3\cdot b, 4\cdot d\}$。求

a) $A\cup B$　　b) $A\cap B$　　c) $A-B$　　d) $B-A$

e) $A+B$

50. 假定 A 是个多重集，其元素是某大学一个系需要的设备的类型，而元素的重数则是该类设备的件数；B 是同一所大学另一个系类似的多重集。例如 A 可以是多重集 {107 · PC, 44 · routers, 6 · servers}，而 B 可以是 {14 · PC, 6 · routers, 2 · mainframes}，其中 routers 为路由器，servers 为服务器，而 mainframes 为大型计算机。

a) 假定两个系使用同一个设备，A 和 B 的什么组合代表该大学应该买的设备？

b) 假定两个系使用同一个设备，A 和 B 的什么组合代表两个系都使用的设备？

c) 假定两个系使用同一个设备，A 和 B 的什么组合代表第二个系使用，但第一个系不使用的设备？

d) 假定两个系不共享设备，A 和 B 的什么组合代表该大学应该购买的设备？

人工智能中使用模糊集合。全集 U 中每个元素在模糊集合 S 中都有个成员度，即 0 和 1 之间(包括 0 和 1)的实数。模糊集合 S 的表示法是列出其元素及成员度(成员度为 0 的元素不列)。例如，用{0.6 Alice，0.9 Brian，0.4 Fred，0.1 Oscar，0.5 Rita}表示名人集合 F，说明 F 中 Alice 的成员度为 0.6，Brian 为 0.9，Fred 为 0.4，Oscar 为 0.1，而 Rita 为 0.5(因此这些人里面 Brian 最出名而 Oscar 最不出名)。再假定 R 是富人集合，$R=${0.4 Alice，0.8 Brian，0.2 Fred，0.9 Oscar，0.7 Rita}。

51. 模糊集合 S 的补是集合 $\overline{S}$，元素在 $\overline{S}$ 中的成员度等于 1 减去该元素在 S 中的成员度。求 $\overline{F}$(不出名者的模糊集合)和 $\overline{S}$(不富裕者的模糊集合)。
52. 模糊集合 S 和 T 的并集是模糊集合 $S\cup T$，其中每个元素的成员度是该元素在 S 和 T 中成员度的最大值。求名人或富人的模糊集合 $F\cup R$。
53. 模糊集合 S 和 T 的交集是模糊集合 $S\cap T$，其中每个元素的成员度是该元素在 S 和 T 中的成员度的最小值。求既出名又富裕者的模糊集合 $F\cap R$。

1.8 函数

1.8.1 引言

在许多情况下我们都会为一个集合的每个元素指派另一个集合(可以就是第一个集合)的一个特定元素。例如，假定对离散数学课的每个学生从{A，B，C，D，F}中选一个字母作为他的得分。再假定 Adams 的得分是 A，Chou 的得分是 C，Goodfriend 的得分是 B，Rodriguez 的得分是 A，而 Stevens 的得分是 F。这一打分如图 1-8 所示。

这种打分就是一个函数。离散数学中函数的概念分外重要。在定义像序列和字符串这样的离散结构时用到函数。函数还用于表示解决一定规模的问题需要的计算机机时。借助自身定义的函数称为递归函数。计算机科学处处使用的这种函数将在第 3 章再讨论。这一节只介绍与离散数学有关的函数的基本概念。

定义 1 令 A 和 B 为集合。从 A 到 B 的函数 f 是对元素的一种指派，对 A 的每个元素恰好指派的 B 的一个元素。如果 f 指派给 A 中元素 a 的唯一的 B 元素是 b，就写成 $f(a)=b$。如果 f 是从 A 到 B 的函数，就写成 $f: A\to B$。

有许多描述函数的方式。有时明确说明指派关系，但常用的方式是给出指派公式来定义函数，如 $f(x)=x+1$。有时也用计算机程序来描述函数。

注意 函数 f: $A\to B$ 有时定义为从 A 到 B 的关系。见 1.6 节。我们将在第 7 章讨论这个方法。

定义 2 如果 f 是从 A 到 B 的函数，就说 A 是 f 的定义域，而 B 是 f 的伴域。如果 $f(a)=b$，就说 b 是 a 的像而 a 是 b 的原像。A 中元素的所有像元素的集合称为 f 的值域。若 f 是从 A 到 B 的函数，有时也说成 f 把 A 映射到 B。

图 1-9 表示一个从 A 到 B 的函数。

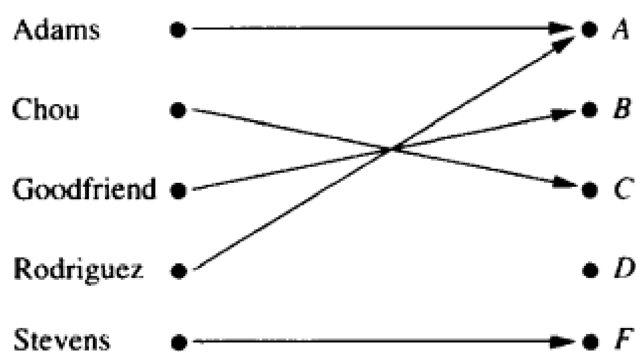


图 1-8 离散数学课成绩

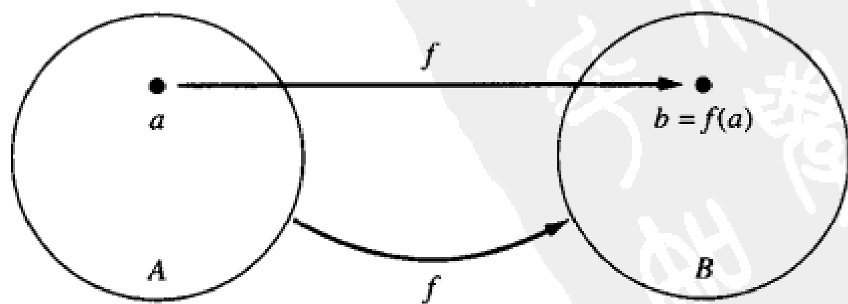


图 1-9 函数 f 把 A 映射到 B

考虑本节开头的例子。令 G 为给离散数学课每个学生打分的函数。例如，$G(\text{Adams})=A$。G 的定义域是集合{Adams，Chou，Goodfriend，Rodriguez，Stevens}，而伴域则是集合$\{A, B, C, D, F\}$。G 的值域是集合$\{A, B, C, F\}$，因为除 D 以外的分值都有学生获得。再考虑下面几个例子。

例 1 令 f 为这样的函数，它把长度为 2 或大于 2 的位串的最后两个字位指派给该位串。于是 f 的定义域是长度不小于 2 的所有位串，而伴域和值域均为集合{00，01，10，11}。 ■

例 2 令 f 为从 **Z** 到 **Z** 的函数，它指派给每个整数的是该整数的平方。于是 $f(x)=x^2$，而 f 的定义域是所有整数的集合，f 的伴域可以从所有整数的集合中选择，f 的值域是所有非负整数中那些完全平方的集合，即{0，1，4，9，…}。 ■

例 3 （适合于熟悉 Pascal 的学生）函数的定义域和伴域往往以程序语言描述。例如 Java 语句 int **floor**(float real)){…}和 Pascal 语句

function *floor*(*x*:**real**):**integer**

都说的是，floor 函数的定义域是实数集合，而它的伴域是整数集合。 ■

具有相同定义域的两个实数值函数可以相加和相乘。

定义 3 令 f_1 和 f_2 是从 A 到 **R** 的函数，那么 f_1+f_2 和 f_1f_2 也是从 A 到 **R** 的函数，其定义为

$$(f_1+f_2)(x)=f_1(x)+f_2(x)$$
$$(f_1f_2)(x)=f_1(x)f_2(x)$$

注意 f_1+f_2 和 f_1f_2 的定义是利用 f_1 和 f_2 在 x 的值计算它们在 x 的值。

例 4 令 f_1 和 f_2 是从 **R** 到 **R** 的函数，且 $f_1(x)=x^2$ 而 $f_2(x)=x-x^2$。函数 f_1+f_2 和 f_1f_2 是什么？

解 从函数的和与积的定义知

$$(f_1+f_2)(x)=f_1(x)+f_2(x)=x^2+(x-x^2)=x$$

而

$$(f_1f_2)(x)=x^2(x-x^2)=x^3-x^4$$ ■

若 f 是从集合 A 到集合 B 的函数，则可以定义 A 的子集的像。

定义 4 令 f 为从集合 A 到集合 B 的函数，S 为 A 的一个子集。S 的像是由 S 中元素的像组成的 B 的子集。我们用 $f(S)$ 表示 S 的像，于是

$$f(S)=\{f(s)\mid s\in S\}$$

例 5 令 $A=\{a, b, c, d, e\}$ 而 $B=\{1, 2, 3, 4\}$，且 $f(a)=2$，$f(b)=1$，$f(c)=4$，$f(d)=1$ 及 $f(e)=1$。子集 $S=\{b, c, d\}$ 的像是集合 $f(S)=\{1, 4\}$。 ■

1.8.2 一对一函数和映上函数

有些函数在它们的定义域的不同成员上有不同的像。这种函数称为一对一的。

定义 5 函数 f 称为一对一的或单射的，当且仅当对于 f 的定义域中的所有 x 和 y，$f(x)=f(y)$ 蕴含着 $x=y$。一对一的函数称为单射。

注意 函数 f 是一对一的，当且仅当只要 $x\neq y$ 就有 $f(x)\neq f(y)$。这种表达 f 为一对一函数的方式是对定义中的蕴含倒置而来。注意我们可以用量词，如 $\forall x\forall y(f(x)=f(y)\rightarrow x=y)$ 或等价地 $\forall x\forall y(x\neq y\rightarrow f(x)\neq f(y))$，来表达 f 是一对一的，其中论域是函数的定义域。

我们通过一对一的函数和不是一对一的函数示例来说明这个概念。

例 6 判断从$\{a, b, c, d\}$到$\{1, 2, 3, 4, 5\}$的函数 f 是否为一对一的，f 的定义是 f

$(a)=4$，$f(b)=5$，$f(c)=1$ 而 $f(d)=3$。

解 f 是一对一的，因为 f 在它定义域的四个元素上取不同的值。图 1-10 说明了这一点。■

例 7 判断从整数集合到整数集合的函数 $f(x)=x^2$ 是否为一对一的。

解 函数 $f(x)=x^2$ 不是一对一的，因为，例如 $f(1)=f(-1)=1$，但 $1\neq-1$。注意若定义域限制为 $\mathbf{Z}^+$，函数 f 就是一对一的。■

例 8 判断函数 $f(x)=x+1$ 是否为一对一的。

解 函数 $f(x)=x+1$ 是一对一的。要证明这一点，只须注意在 $x\neq y$ 时 $x+1\neq y+1$。■

现在我们给出保证函数为一对一的某些条件。

定义 6 定义域和伴域都是实数集合子集的函数 f 称为*严格递增的*，如果对 f 的定义域中的 x 和 y，只要 $x<y$ 就有 $f(x)<f(y)$。类似地，f 称为*严格递减的*，如果对 f 定义域中的 x 和 y，只要 $x<y$ 就有 $f(x)>f(y)$。

> **注意** 如果 $\forall x\forall y((x<y)\rightarrow(f(x)<f(y)))$，函数 f 是严格递增的；如果 $\forall x\forall y((x<y)\rightarrow(f(x)>f(y)))$，函数 f 就是严格递减的。其中论域均为 f 的定义域。

从上述定义可知，只要函数是严格递增的或严格递减的，它必定是一对一的。

有些函数的值域和伴域相等，也就是说伴域中的每个成员都是定义域中某个元素的像。具有这一性质的函数称为映上函数。

定义 7 从 A 到 B 的函数 f 称为*映上的*或*满射的*，当且仅当对每个 $b\in B$，有元素 $a\in A$ 使得 $f(a)=b$。如果函数 f 是映上的，就说它是*满射函数*。

> **注意** 如果 $\forall y\exists x(f(x)=y)$，函数 f 就是映上的，其中 x 的论域是函数的定义域，y 的论域是函数的伴域。

我们现在举几个映上函数和非映上函数的例子。

例 9 令 f 为从 $\{a, b, c, d\}$ 到 $\{1, 2, 3\}$ 的函数，其定义为 $f(a)=3$，$f(b)=2$，$f(c)=1$ 及 $f(d)=3$。f 是映上函数吗？

解 由于伴域中所有 3 个元素均为定义域中元素的像，f 是映上的。图 1-11 说明了这一点。注意若伴域是 $\{1, 2, 3, 4\}$，f 就不是映上的。■

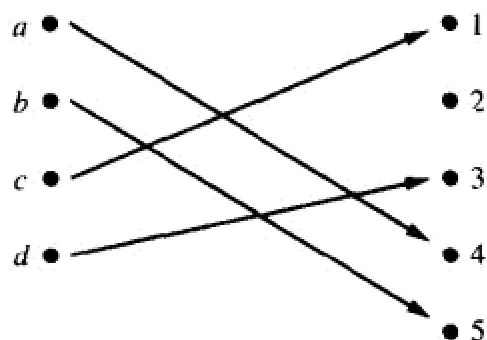


图 1-10 一个一对一函数

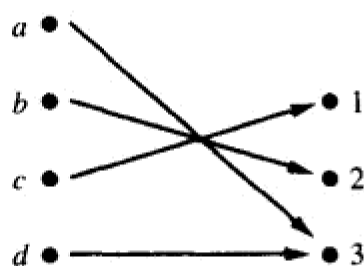


图 1-11 一个映上函数

例 10 从整数集到整数集的函数 $f(x)=x^2$ 是映上的吗？

解 f 不是映上的，因此，比如说没有 x 使 $x^2=-1$。■

例 11 从整数集到整数集的函数 $f(x)=x+1$ 是映上的吗？

解 这个函数是映上的，因为对每个整数 y，都有一个整数 x 使 $f(x)=y$。为看出这一点，只要注意 $f(x)=y$ 的充分必要条件是 $x+1=y$，而这只要令 $x=y-1$ 就成立。■

定义 8 若函数 f 既是一对一的，又是映上的，就说它是*一一对应*或*双射的*。

下面几个例子阐述双射的概念。

例 12 令 f 为从 $\{a, b, c, d\}$ 到 $\{1, 2, 3, 4\}$ 的函数，其定义为 $f(a)=4$，$f(b)=2$，$f(c)=1$ 及 $f(d)=3$。f 是双射吗？

解 函数 f 是一对一的和映上的。它是一对一的，是因为函数值都不同；它是映上的，是因为伴域中的所有 4 个元素，均为定义域的元素的像。于是 f 是双射。■

图 1-12 给出了 4 个函数：其中第 1 个是一对一的，但不是映上的；第 2 个是映上的，但不是一对一的；第 3 个，既是一对一的，又是映上的；第 4 个既不是一对一的，也不是映上的。图 1-12 中的第 5 个对应关系不是函数，因为它把一个元素传递给两个不同的元素。

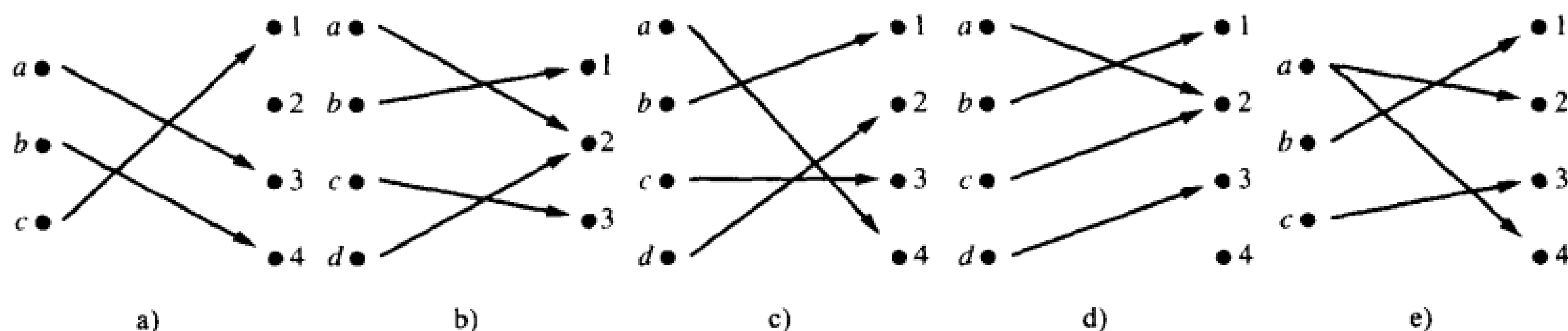


图 1-12 不同类型的对应关系的例子

a)一对一，非映上；b)映上，非一对一；c)一对一，映上；d)既非一对一，也非映上；e)不是函数

假定 f 是从集合 A 到它自己的函数。如果 A 是有限的，那么 f 是一对一的当且仅当它是映上的。（从本节末练习 64 的结果即可得出这一结论。）当 A 为无限集时，这一结论不一定成立（参看 3.2 节）。

例 13 令 A 为集合。A 上的恒等函数是函数 $\iota_A: A \to A$，其中

$$\iota_A(x) = x$$

对所有 $x \in A$。换言之，恒等函数 ι_A 是这样的函数，它赋给每个元素的是这个元素自身。函数 ι_A 是一对一的和映上的，所以是双射。■

1.8.3 反函数和函数组合

现在考虑从集合 A 到集合 B 的一一对应 f。由于 f 是映上的函数，B 的每个元素都是 A 中某元素的像。又由于 f 还是一对一的函数，B 的每个元素都是 A 中唯一的一个元素的像。于是我们可以定义一个从 B 到 A 的新函数，把 f 给出的对应关系颠倒过来。

定义 9 令 f 为从集合 A 到集合 B 的一一对应，f 的反函数是这样的函数，它指派给 B 中元素 b 的是 A 中使得 $f(a)=b$ 唯一元素 a。f 的反函数用 f^{-1} 表示，于是在 $f(a)=b$ 时 $f^{-1}(b)=a$。

图 1-13 给出了反函数概念的说明。

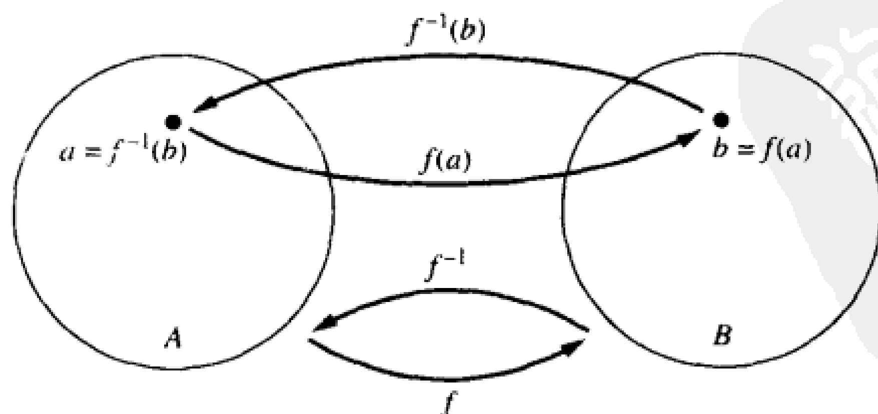


图 1-13 函数 f^{-1} 是函数 f 的反函数

如果函数 f 不是一一对应，就无法定义反函数。若 f 不是一一对应，那么它或者不是一对

一的，或者不是映上的，如果 f 不是一对一的，则伴域中的某元素 b 是定义域中多个元素的像。如果 f 不是映上的，那么伴域中某个元素 b 不是定义域中任何元素的像，即定义域中不存在元素 a 使 $f(a)=b$。总之如果 f 不是一一对应，就不能为伴域中每个元素 b，都指派定义域中唯一的元素 a，使 $f(a)=b$(因为对某个 b 或者有多个这样的 a，或者没有这样的 a)。

一一对应关系称为可逆的，因为作为函数可以定义它的反函数。如果函数不是一一对应关系，就说它是不可逆的，因为这样的函数没有反函数。

例 14 令 f 为从 $\{a, b, c\}$ 到 $\{1, 2, 3\}$ 的函数，使 $f(a)=2$，$f(b)=3$ 及 $f(c)=1$。f 可逆吗？如果可逆，其反函数是什么？

解 f 是可逆的，因为它是一个一对一的对应关系。其反函数 f^{-1} 颠倒 f 给出的对应关系，所以 $f^{-1}(1)=c$，$f^{-1}(2)=a$ 而 $f^{-1}(3)=b$。■

例 15 令 f 为从整数集到整数集的函数，使得 $f(x)=x+1$。f 可逆吗？如果可逆，其反函数是什么？

解 f 可逆，因为前面已证明它是一一对应关系。要颠倒对应关系，设 y 是 x 的像，即 $y=x+1$。从而 $x=y-1$。即 $y-1$ 是 $\mathbf{Z}$ 的唯一元素，在 f 之下与 y 对应，于是 $f^{-1}(y)=y-1$。■

例 16 令 f 是从 $\mathbf{Z}$ 到 $\mathbf{Z}$ 的函数，使 $f(x)=x^2$。f 可逆吗？

解 由于 $f(-1)=f(1)=1$，f 不是一对一的，要想定义反函数，就得为 1 指派两个元素。因此 f 是不可逆的。■

定义 10 令 g 为从集合 A 到集合 B 的函数，f 是从集合 B 到集合 C 的函数，函数 f 和 g 的组合用 $f\circ g$ 表示，定义为

$$(f\circ g)(a) = f(g(a))$$

换句话说，函数 $f\circ g$ 指派给 A 的元素 a 的就是 f 指派给 $g(a)$ 的。注意，如果 g 的值域不是 f 的定义域的子集，就无法定义 $f\circ g$，图 1-14 解释了函数组合的概念。

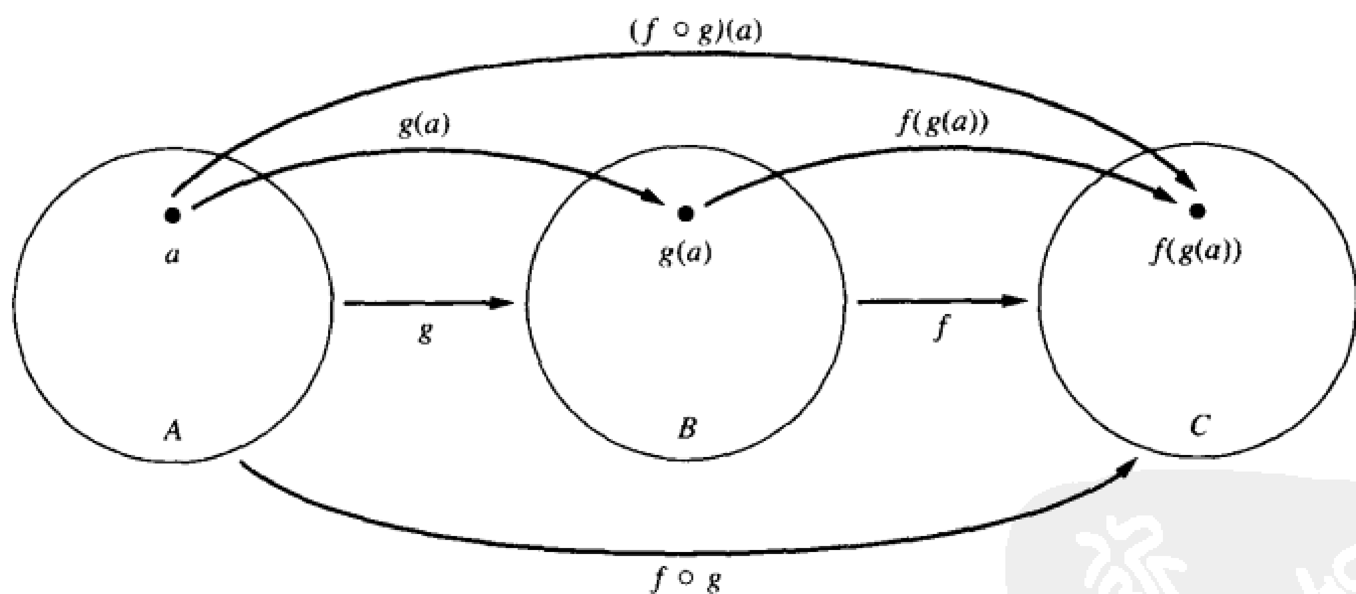


图 1-14 函数 f 和 g 的组合

例 17 令 g 为从 $\{a, b, c\}$ 到它自己的函数，使 $g(a)=b$，$g(b)=c$，而 $g(c)=a$。令 f 为从 $\{a, b, c\}$ 到 $\{1, 2, 3\}$ 的函数，使 $f(a)=3$，$f(b)=2$，而 $f(c)=1$。f 和 g 的组合是什么？g 和 f 的组合是什么？

解 $f\circ g$ 的定义是 $(f\circ g)(a)=f(g(a))=f(b)=2$，$(f\circ g)(b)=f(g(b))=f(c)=1$，而 $(f\circ g)(c)=f(g(c))=f(a)=3$。

注意 $g\circ f$ 是没有定义的，因为 f 的值域不是 g 的定义域的一部分。■

例 18 令 f 和 g 为从整数集到整数集的函数，其定义为 $f(x)=2x+3$ 和 $g(x)=3x+2$。f 和 g 的组合是什么？g 和 f 的组合是什么？

解 $f\circ g$ 和 $g\circ f$ 均有定义，即

$$(f\circ g)(x)=f(g(x))=f(3x+2)=2(3x+2)+3=6x+7$$

及

$$(g\circ f)(x)=g(f(x))=g(2x+3)=3(2x+3)+2=6x+11$$ ■

注意 尽管例 18 中对函数 f 和 g 而言 $f\circ g$ 和 $g\circ f$ 均有定义，$f\circ g$ 和 $g\circ f$ 并不相等。换言之，对函数组合而言交换律不成立。

在构造函数和它的反函数的组合时，不论以什么次序组合，得到的都是恒等函数。要看清这一点，假定 f 是从集合 A 到集合 B 的一一对应关系，那么存在反函数 f^{-1}，而且 f^{-1} 是从 B 到 A 的一一对应关系。反函数把原函数的对应关系颠倒过来，所以当 $f(a)=b$ 时 $f^{-1}(b)=a$，当 $f^{-1}(b)=a$ 时，$f(a)=b$。因此，

$$(f^{-1}\circ f)(a)=f^{-1}(f(a))=f^{-1}(b)=a$$

及

$$(f\circ f^{-1})(b)=f(f^{-1}(b))=f(a)=b$$

从而 $f^{-1}\circ f=\iota_A$，而 $f\circ f^{-1}=\iota_B$，其中 ι_A 和 ι_B 分别是集合 A 和 B 上的恒等函数。这就是说 $(f^{-1})^{-1}=f$。

1.8.4 函数的图像

从 A 到 B 的每个函数都对应着 $A\times B$ 中的一个有序偶集合，这个序偶集合称为该函数的图像，并且往往用图来表示以帮助理解函数的行为。

定义 11 令 f 为从集合 A 到集合 B 的函数，函数 f 的图像是序偶集合 $\{(a,b)\mid a\in A$ 且 $f(a)=b\}$。

根据定义，从 A 到 B 的函数的图像是 $A\times B$ 的子集，这个子集包含的序偶中第二项等于 B 中由 f 指派给第一项的那个元素。

例 19 图示从整数集到整数集的函数 $f(n)=2n+1$ 的图像。

解 f 的图像是形为 $(n, 2n+1)$ 的序偶的集合，其中 n 为整数。图 1-15 是其图像显示。 ■

例 20 图示整数集到整数集的函数 $f(x)=x^2$。

解 f 的图像是形为 $(x, f(x))=(x, x^2)$ 的序偶的集合，其中 x 为整数。图1-16 是其图像显示。 ■

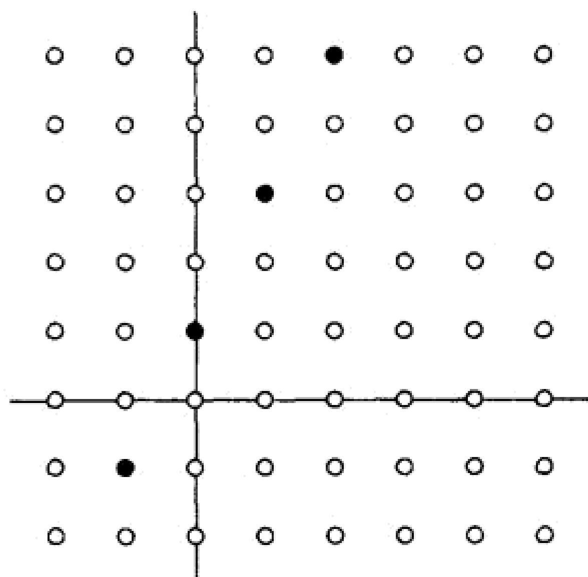


图 1-15 从 **Z** 到 **Z** 的函数 $f(n)=2n+1$ 的图像

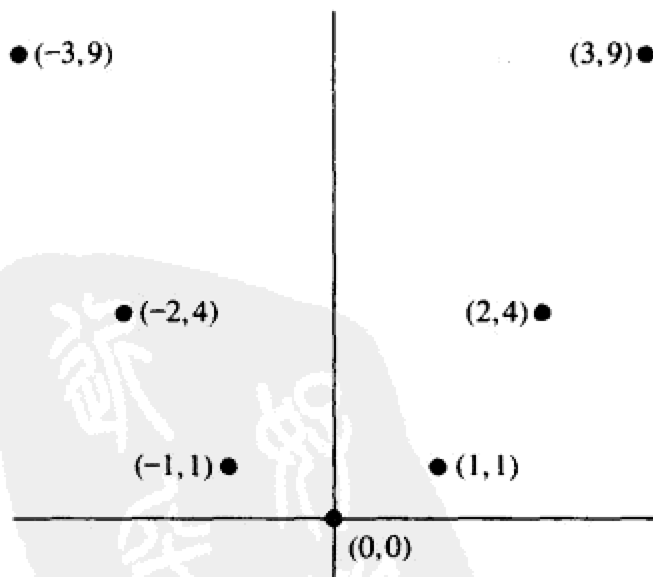


图 1-16 从 **Z** 到 **Z** 的 $f(x)=x^2$ 的图像

1.8.5 几个重要的函数

下面介绍离散数学中两个重要的函数，即下取整函数和上取整函数。令 x 为实数。下取整函数把 x 下舍入到小于或等于 x 又最接近 x 的整数，而上取整函数则把 x 上舍入到大于或等于 x 又最接近 x 的整数。在统计对象个数时常使用这两个函数。在分析解一定规模的问题的计算机过程使用的步数时，这两个函数起着重要的作用。

定义 12 下取整函数指派给实数 x 的是小于或等于 x 的最大整数。下取整函数在 x 的值用 $\lfloor x \rfloor$ 表示。上取整函数指派给实数 x 的是大于或等于 x 的最小整数。上取整函数在 x 的值用 $\lceil x \rceil$ 表示。

注意 下取整函数也常称为最大整数函数，这时往往用 $[x]$ 表示。

例 21 下面是下取整函数和上取整函数的若干值：

$$\lfloor 1/2 \rfloor = 0, \quad \lceil 1/2 \rceil = 1, \quad \lfloor -1/2 \rfloor = -1, \quad \lceil -1/2 \rceil = 0,$$
$$\lfloor 3.1 \rfloor = 3, \quad \lceil 3.1 \rceil = 4, \quad \lfloor 7 \rfloor = 7, \quad \lceil 7 \rceil = 7$$ ■

图 1-17 显示的是下取整函数和上取整函数的图像。在图 1-17a 中，显示下取整函数 $\lfloor x \rfloor$。这个函数在 $[n, n+1)$ 区间取同样的值 n，然后，当 $x=n+1$ 时，取值跳到 $n+1$。在图 1-17b 中，显示上取整函数 $\lceil x \rceil$。这个函数在 $(n, n+1]$ 区间取同样的值 $n+1$，然后，当 x 取略大于 $n+1$ 时，取值跳到 $n+2$。

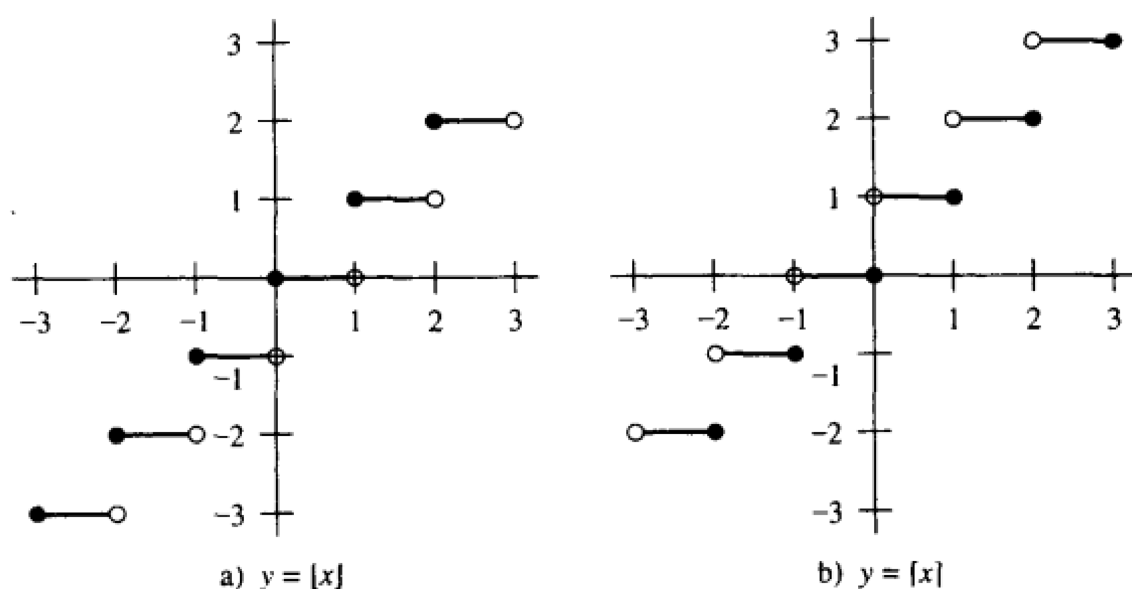


图 1-17 a)下取整函数图像 b)上取整函数图像

下取整函数和上取整函数有广泛的应用，包括用于数据存储和数据传输。考虑例 22 和例 23，这是研究数据库和数据通信问题时要完成的典型的基本计算。

例 22 存在计算机磁盘上的数据或数据网络上传输的数据通常表示为字节串。每个字节由 8 个字位组成，要表示 100 字位的数据需要多少字节？

解 要决定需要的字节数，就要找出最小的整数，它至少要与 100 除以 8 的商一样大，8 是每个字节的字位数。于是，需要的字节数是 $\lceil 100/8 \rceil = \lceil 12.5 \rceil = 13$。 ■

例 23 异步传输模式(ATM)(用于骨干网络上的通信协议)下，数据按 53 个字节分组，每组称为一个信元。以速率每秒 500 千字位传输数据的连接上一分钟能传输多少个 ATM 信元？

解 一分钟内这一连接能传输 $500\,000 \times 60 = 30\,000\,000$ 字位。每个 ATM 信元长度是 53 字节，也就是 $53 \times 8 = 424$ 字位。要计算一分钟能传输多少信元，需计算不超过 30 000 000 除以 424 的最大整数。于是在每秒 500 千字位的连接上一分钟能传输的 ATM 信元数是 $\lfloor 30\,000\,000/424 \rfloor = 70\,754$。 ■

以 x 代表实数的表 1-23 给出了下取整函数和上取整函数的若干简单但重要的性质。由于这两个函数在离散数学中出现得十分频繁，看一看表中的恒等式是有益的。表中的每条性质都可

以用下取整函数和上取整函数的定义来建立。性质(1a)，(1b)，(1c)和(1d)可以直接由定义得出。例如，(1a)说的是$\lfloor x\rfloor=n$当且仅当n小于或等于x而$n+1$大于x。这恰恰就是n为不超过x的最大整数的含义，也就是$\lfloor x\rfloor=n$的定义。性质(1b)，(1c)和(1d)可以类似地建立起来。我们来证明(4a)是成立的。

表 1-23　上取整函数和下取整函数的有用性质(n为整数)

(1a) $\lfloor x\rfloor=n$ 当且仅当 $n\leqslant x<n+1$ (1b) $\lceil x\rceil=n$ 当且仅当 $n-1<x\leqslant n$ (1c) $\lfloor x\rfloor=n$ 当且仅当 $x-1<n\leqslant x$ (1d) $\lceil x\rceil=n$ 当且仅当 $x\leqslant n<x+1$
(2) $x-1<\lfloor x\rfloor\leqslant x\leqslant\lceil x\rceil<x+1$
(3a) $\lfloor -x\rfloor=-\lceil x\rceil$ (3b) $\lceil -x\rceil=-\lfloor x\rfloor$
(4a) $\lfloor x+n\rfloor=\lfloor x\rfloor+n$ (4b) $\lceil x+n\rceil=\lceil x\rceil+n$

证　假定$\lfloor x\rfloor=m$，其中m为整数。由(1a)知$m\leqslant x<m+1$。在这一不等式中加上n得$m+n\leqslant x+n<m+n+1$。再次利用(1a)，可知$\lfloor x+n\rfloor=m+n=\lfloor x\rfloor+n$，这正是我们要证明的。其他性质的证明推迟到练习中去做。■

除了表1-23列出的性质，上取整函数和下取整函数还有许多其他有用的性质。也有许多关于这些函数的语句看似正确，实则不然。我们将在例24和例25中考虑与上取整函数和下取整函数有关的语句。

在考虑下取整函数相关的语句时，一个有用的方法是令$x=n+\epsilon$，其中$n=\lfloor x\rfloor$是一个整数，ϵ是x的分数部分，满足不等式$0\leqslant\epsilon<1$。类似地，考虑上取整函数时，通常写$x=n-\epsilon$，其中$n=\lceil x\rceil$且$0\leqslant\epsilon<1$。

例 24　证明若x是一个实数，那么$\lfloor 2x\rfloor=\lfloor x\rfloor+\left\lfloor x+\frac{1}{2}\right\rfloor$。

解　要证明这个语句，令$x=n+\epsilon$，其中n是正整数且$0\leqslant\epsilon<1$。依据ϵ是否小于$\frac{1}{2}$，有两种情况要考虑。(选择这两种情况的原因见下面的证明。)

首先考虑$0\leqslant\epsilon<\frac{1}{2}$的情况。此时，$2x=2n+2\epsilon$且$\lfloor 2x\rfloor=2n$，因为$0\leqslant 2\epsilon<1$。类似地，$x+\frac{1}{2}=n+\left(\frac{1}{2}+\epsilon\right)$。使得$\left\lfloor x+\frac{1}{2}\right\rfloor=n$，因为$0<\frac{1}{2}+\epsilon<1$。因此，$\lfloor 2x\rfloor=2n$且$\lfloor x\rfloor+\left\lfloor x+\frac{1}{2}\right\rfloor=n+n=2n$。

接下来，考虑$\frac{1}{2}\leqslant\epsilon<1$的情况。此时，$2x=2n+2\epsilon=(2n+1)+(2\epsilon-1)$。由于$0\leqslant 2\epsilon-1<1$，可得$\lfloor 2x\rfloor=2n+1$。因为$\left\lfloor x+\frac{1}{2}\right\rfloor=\left\lfloor n+\left(\frac{1}{2}+\epsilon\right)\right\rfloor=\left\lfloor n+1+\left(\epsilon-\frac{1}{2}\right)\right\rfloor$且$0\leqslant\epsilon-\frac{1}{2}<1$，可得$\left\lfloor x+\frac{1}{2}\right\rfloor=n+1$。因此，$\lfloor 2x\rfloor=2n+1$且$\lfloor x\rfloor+\left\lfloor x+\frac{1}{2}\right\rfloor=n+(n+1)=2n+1$。证毕。■

例 25　对全体实数x和y，证明或否定$\lceil x+y\rceil=\lceil x\rceil+\lceil y\rceil$。

解　尽管这个语句看似合理，它其实是假的。一个反例就是令$x=\frac{1}{2}$且$y=\frac{1}{2}$，此时

$$\lceil x+y\rceil=\left\lceil\frac{1}{2}+\frac{1}{2}\right\rceil=\lceil 1\rceil=1\text{，但}\lceil x\rceil+\lceil y\rceil=\left\lceil\frac{1}{2}\right\rceil+\left\lceil\frac{1}{2}\right\rceil=1+1=2\text{。}$$ ■

有几类函数是本课本一直要用到的，其中包括多项式、对数和指数函数。附录 1 给出了课本中需要的这些函数之性质的简要回顾。本书中用记号 $\log x$ 表示 x 以 2 为基的对数，因为 2 是我们将经常使用的对数基数。我们用 $\log_b x$ 表示以 b 为基的对数，其中 b 是大于 1 的任意实数。用 $\ln x$ 表示自然对数。

另一种我们将在本书中使用的函数是*阶乘函数* f: $\mathbf{N}\to\mathbf{Z}^+$，表示为 $f(n)=n!$。$f(n)=n!$ 的值是前几个正整数的积，因此 $f(n)=1\cdot 2\cdots(n-1)\cdot n[f(0)=0!=1]$。

例 26 $f(6)=6!=1\times2\times3\times4\times5\times6=720$。 ■

练习

1. 为什么下列几问中的 f 不是从 **R** 到 **R** 的函数？
   a) $f(x)=1/x$　　b) $f(x)=\sqrt{x}$　　c) $f(x)=\pm\sqrt{(x^2+1)}$
2. 判断下面定义的几个 f 是不是从 **Z** 到 **R** 的函数。
   a) $f(n)=\pm n$　　b) $f(n)=\sqrt{n^2+1}$　　c) $f(n)=1/(n^2-4)$
3. 判断 f 是否为从所有位串的集合到整数集合的函数：
   a) $f(S)$是 S 中某个 0 字位的位置。
   b) $f(S)$是 S 中 1 字位的个数。
   c) $f(S)$是最小整数 i 使 S 中的第 i 字位为 1，当 S 是不含字位的空串时 $f(S)=0$。
4. 求下列函数的定义域和值域。
   a)函数在每个非负整数的值为该整数的最后一个数字。
   b)函数在每个正整数的值是比它小的最大整数。
   c)函数在每个位串的值是串中 1 字位的个数。
   d)函数在每个位串的值是串的字位数。
5. 求下列函数的定义域和值域。
   a)函数在每个位串的值是串中 1 字位个数与 0 字位个数之差。
   b)函数在每个位串的值是串中 0 字位个数的两倍。
   c)函数在每个位串的值是把串分成字节(8 个字位为一字节)时不够一个字节的字位数。
   d)函数在每个正整数的值是不超过此整数的最大完全平方。
6. 求下列函数的定义域和值域。
   a)函数在每对正整数的值是对中的第一个整数。
   b)函数在每个正整数的值为该整数的最大十进位。
   c)函数在一个位串的值是串中 1 字位个数与 0 字位个数之差。
   d)函数在每个正整数的值是不超过该整数的平方根的最大的整数。
   e)函数在一个位串的值是串中最长的 1 字位串。
7. 求下列函数的定义域和值域。
   a)函数在每对正整数的值是这两个整数中的最大的整数。
   b)函数在每个正整数的值是 0，1，2，3，4，5，6，7，8，9 中未在整数数位中出现的数。
   c)函数在一个位串的值是串中块 11 出现的次数。
   d)函数在一个位串的值是串中第一个 1 的数值位置，若位串为全 0，就赋值 0。
8. 求下列各值：
   a) $\lfloor 1.1\rfloor$　　b) $\lceil 1.1\rceil$　　c) $\lfloor -0.1\rfloor$　　d) $\lceil -0.1\rceil$
   e) $\lceil 2.99\rceil$　　f) $\lceil -2.99\rceil$　　g) $\lfloor 1/2+\lceil 1/2\rceil\rfloor$　　h) $\lceil\lfloor 1/2\rfloor+\lceil 1/2\rceil+1/2\rceil$
9. 求下列各值：
   a) $\lceil 3/4\rceil$　　b) $\lfloor 7/8\rfloor$　　c) $\lceil -3/4\rceil$　　d) $\lfloor -7/8\rfloor$

e) $\lceil 3 \rceil$　　f) $\lfloor -1 \rfloor$　　g) $\lfloor 1/2+\lceil 3/2\rceil \rfloor$　　h) $\lfloor 1/2 \cdot \lfloor 5/2\rfloor \rfloor$

10. 判断下列从$\{a, b, c, d\}$到它自身的函数是否一对一。
   a) $f(a)=b$, $f(b)=a$, $f(c)=c$, $f(d)=d$
   b) $f(a)=b$, $f(b)=b$, $f(c)=d$, $f(d)=c$
   c) $f(a)=d$, $f(b)=b$, $f(c)=c$, $f(d)=d$
11. 练习 10 中哪些函数是映上的?
12. 判断下列 **Z** 到 **Z** 的函数是否一对一。
   a) $f(n)=n-1$　　b) $f(n)=n^2+1$　　c) $f(n)=n^3$　　d) $f(n)=\lceil n/2\rceil$
13. 上一题中哪些函数是映上的?
14. 判断在下列情况下 f: $\mathbf{Z}\times\mathbf{Z}\rightarrow\mathbf{Z}$ 是否映上。
   a) $f(m, n)=2m-n$.　　b) $f(m, n)=m^2-n^2$.
   c) $f(m, n)=m+n+1$.　　d) $f(m, n)=|m|-|n|$.
   e) $f(m, n)=m^2-4$.
15. 判断在下列情况下函数 f: $\mathbf{Z}\times\mathbf{Z}\rightarrow\mathbf{Z}$ 是否是映上的?
   a) $f(m, n)=m+n$.　　b) $f(m, n)=m^2+n^2$.
   c) $f(m, n)=m$.　　d) $f(m, n)=|n|$.
   e) $f(m, n)=m-n$.
16. 给出从 **N** 到 **N** 的函数的例子，使得:
   a)一对一但非映上。　　b)映上但不一对一。
   c)既映上又一对一(但不同于恒等函数)。　　d)既非映上又非一对一。
17. 给出从整数集合到正整数集合的函数的明确公式，使得:
   a)一对一但非映上。　　b)映上但非一对一。
   c)既映上又一对一。　　d)既不映上又不一对一。
18. 判断下列各函数是否从 **R** 到 **R** 的双射。
   a) $f(x)=-3x+4$　　b) $f(x)=-3x^2+7$
   c) $f(x)=(x+1)/(x+2)$　　d) $f(x)=x^5+1$
19. 判断下列各函数是否从 **R** 到 **R** 的双射。
   a) $f(x)=2x+1$　　b) $f(x)=x^2+1$
   c) $f(x)=x^3$　　d) $f(x)=(x^2+1)/(x^2+2)$
20. 令 f: $\mathbf{R}\rightarrow\mathbf{R}$ 且 $f(x)>0(x\in\mathbf{R})$。证明 $f(x)$严格递增当且仅当函数 $g(x)=1/f(x)$严格递减。
21. 令 f: $\mathbf{R}\rightarrow\mathbf{R}$ 且 $f(x)>0$。证明 $f(x)$严格递减当且仅当函数 $g(x)=1/f(x)$严格递增。
22. 令 $S=\{-1, 0, 2, 4, 7\}$。求 $f(S)$，如果
   a) $f(x)=1$　　b) $f(x)=2x+1$
   c) $f(x)=\lceil x/5\rceil$　　d) $f(x)=\lfloor (x^2+1)/3\rfloor$
23. 令 $f(x)=\lfloor x^2/3\rfloor$。求 $f(S)$，如果
   a) $S=\{-2, -1, 0, 1, 2, 3\}$　　b) $S=\{0, 1, 2, 3, 4, 5\}$
   c) $S=\{1, 5, 7, 11\}$　　d) $S=\{2, 6, 10, 14\}$
24. 令 $f(x)=2x$。什么是
   a) $f(\mathbf{Z})$?　　b) $f(\mathbf{N})$?　　c) $f(\mathbf{R})$?
25. 假定 g 是从 A 到 B 的函数，f 是从 B 到 C 的函数。
   a)证明如果 f 和 g 均为一对一函数，那么 $f\circ g$ 也是。
   b)证明如果 f 和 g 均为到映上函数，那么 $f\circ g$ 也是。

*26. 如果 f 和 $f\circ g$ 都是一对一的，能否得出结论 g 也是一对一的? 说明理由。
*27. 如果 f 和 $f\circ g$ 都是映上的，能否得出结论 g 也是映上的? 说明理由。

28. 已知 $f(x)=x^2+1$ 和 $g(x)=x+2$ 都是从 **R** 到 **R** 的函数，求 $f\circ g$ 和 $g\circ f$。
29. 对上题中的函数 f 和 g 求 $f+g$ 和 fg。

30. 令 $f(x)=ax+b$，$g(x)=cx+d$，其中 a，b，c，d 为常数。哪些常数 a，b，c，d 能使 $f\circ g=g\circ f$？
31. 求证从 $\mathbf{R}$ 到 $\mathbf{R}$ 的函数 $f(x)=ax+b$ 是可逆的，其中 a 和 b 为常数且 $a\neq 0$。找出 f 的反函数。
32. 令 f 为从集合 A 到集合 B 的函数，令 S 和 T 为 A 的子集，求证
    a) $f(S\cup T)=f(S)\cup f(T)$　　b) $f(S\cap T)\subseteq f(S)\cap f(T)$
33. 给出一个例子说明上题 b)中可能是真子集。

令 f 为从集合 A 到集合 B 的函数，S 为 B 的子集。S 的逆像定义为 A 的子集，该子集包含 S 的所有元素的原像。我们用 $f^{-1}(S)$ 表示 S 的逆像，于是 $f^{-1}(S)=\{a\in A\mid f(a)\in S\}$。

34. 令 f 为从 $\mathbf{R}$ 到 $\mathbf{R}$ 的函数 $f(x)=x^2$。求
    a) $f^{-1}(\{1\})$　　b) $f^{-1}(\{x\mid 0<x<1\})$
    c) $f^{-1}(\{x\mid x>4\})$
35. 令 $g(x)=\lfloor x\rfloor$。求
    a) $g^{-1}(\{0\})$　　b) $g^{-1}(\{-1,0,1\})$
    c) $g^{-1}(\{x\mid 0<x<1\})$
36. 令 f 为从 A 到 B 的函数，令 S 和 T 为 B 的子集，求证
    a) $f^{-1}(S\cup T)=f^{-1}(S)\cup f^{-1}(T)$　　b) $f^{-1}(S\cap T)=f^{-1}(S)\cap f^{-1}(T)$
37. 令 f 为从 A 到 B 的函数，S 为 B 的子集。求证 $f^{-1}(\overline{S})=\overline{f^{-1}(S)}$。
38. 证明 $\lfloor x+1/2\rfloor$ 是最接近 x 的整数，除非 x 恰为两个(相邻)整数的中间数，此时它为这两个整数中较大的一个。
39. 证明 $\lceil x-1/2\rceil$ 是最接近 x 的整数，除非 x 恰为两个(相邻)整数的中间数，此时，它为这两个整数中较小的一个。
40. 证明对实数 x，当 x 不是整数时 $\lceil x\rceil-\lfloor x\rfloor=1$，而在 x 为整数时 $\lceil x\rceil-\lfloor x\rfloor=0$。
41. 证明对实数 x，$x-1<\lfloor x\rfloor\leqslant x\leqslant\lceil x\rceil<x+1$。
42. 若 x 为实数，m 为整数，求证 $\lceil x+m\rceil=\lceil x\rceil+m$。
43. 若 x 为实数，n 为整数，求证
    a) $x<n$ 当且仅当 $\lfloor x\rfloor<n$。　　b) $n<x$ 当且仅当 $n<\lceil x\rceil$。
44. 若 x 为实数，n 为整数，求证
    a) $x\leqslant n$ 当且仅当 $\lceil x\rceil\leqslant n$。　　b) $n\leqslant x$ 当且仅当 $n\leqslant\lfloor x\rfloor$。
45. 证明若 n 为整数，则当 n 为偶数时 $\lfloor n/2\rfloor=n/2$，n 为奇数时 $\lfloor n/2\rfloor=(n-1)/2$。
46. 证明若 x 为实数，则 $\lfloor -x\rfloor=-\lceil x\rceil$，$\lceil -x\rceil=-\lfloor x\rfloor$。
47. 有些计算器上有个 INT 函数，当 x 为非负实数时 $\mathrm{INT}(x)=\lfloor x\rfloor$，当 x 为负实数时 $\mathrm{INT}(x)=\lceil x\rceil$。求证这一函数 INT 满足等式 $\mathrm{INT}(-x)=-\mathrm{INT}(x)$。
48. 令 a 和 b 为实数，且 $a<b$。用下取整函数及/或上取整函数表示使 $a\leqslant n\leqslant b$ 的整数 n 的数目。
49. 令 a 和 b 的实数，且 $a<b$，用下取整函数及/或上取整函数表示使 $a<n<b$ 的整数 n 的数目。
50. 含 n 个字位的数据需要用几个字节来编码？其中 n 是
    a) 4　　b) 10　　c) 500　　d) 3000
51. 含 n 个字位的数据需要用几个字节编码？其中 n 为
    a) 7　　b) 17　　c) 1001　　d) 28800
52. 用下面给出的传输率的连接在 10 秒钟内能传输多少 ATM 信元(参看例 23)？
    a) 每秒 128 千字位。　　b) 每秒 300 千字位。　　c) 每秒兆字位。
53. 数据在某以太网上以 1500 个八位字节为信息块传输。下面的数据量在这个以太网上各需要多少个信息块传输(注意字节就是 8 位字节)？
    a) 150 千字节的数据。　　b) 384 千字节的数据。
    c) 1.544 兆字节的数据。　　d) 45.3 兆字节的数据。
54. 画出从 $\mathbf{Z}$ 到 $\mathbf{Z}$ 的函数 $f(n)=1-n^2$ 的图像。
55. 画出从 $\mathbf{R}$ 到 $\mathbf{R}$ 的函数 $f(x)=\lfloor 2x\rfloor$ 的图像。

56. 画出从 **R** 到 **R** 的函数 $f(x)=\lfloor x/2\rfloor$的图像。
57. 画出从 **R** 到 **R** 的函数 $f(x)=\lfloor x\rfloor+\lfloor x/2\rfloor$的图像。
58. 画出从 **R** 到 **R** 的函数 $f(x)=\lceil x\rceil+\lfloor x/2\rfloor$的图像。
59. 画出下列各函数的图像：
   a) $f(x)=\lfloor x+1/2\rfloor$　　b) $f(x)=\lfloor 2x+1\rfloor$
   c) $f(x)=\lceil x/3\rceil$　　d) $f(x)=\lceil 1/x\rceil$
   e) $f(x)=\lceil x-2\rceil+\lfloor x+2\rfloor$　　f) $f(x)=\lfloor 2x\rfloor\lceil x/2\rceil$
   g) $f(x)=\lceil\lfloor x-1/2\rfloor+1/2\rceil$
60. 画出下列各函数的图像：
   a) $f(x)=\lceil 3x-2\rceil$　　b) $f(x)=\lceil 0.2x\rceil$
   c) $f(x)=\lfloor -1/x\rfloor$　　d) $f(x)=\lfloor x^2\rfloor$
   e) $f(x)=\lceil x/2\rceil\lfloor x/2\rfloor$　　f) $f(x)=\lfloor x/2\rfloor+\lceil x/2\rceil$
   g) $f(x)=\lfloor 2\lceil x/2\rceil+1/2\rfloor$
61. 求 $f(x)=x^3+1$ 的反函数。
62. 假定 f 是从 Y 到 Z 的可逆函数，g 是从 X 到 Y 的可逆函数。证明函数组合 $f\circ g$ 的反函数$(f\circ g)^{-1}=g^{-1}\circ f^{-1}$。
63. 令 S 为全集 U 的子集。S 的*特征函数* f_S 是从 U 到集合$\{0, 1\}$的函数，使得若 x 属于 S 则 $f_S(x)=1$，若 x 不属于 S，则 $f_S(x)=0$。令 A，B 为集合。求证
   a) $f_{A\cap B}(x)=f_A(x)\cdot f_B(x)$　　b) $f_{A\cup B}(x)=f_A(x)+f_B(x)-f_A(x)\cdot f_B(x)$
   c) $f_{\overline{A}}(x)=1-f_A(x)$　　d) $f_{A\oplus B}(x)=f_A(x)+f_B(x)-2f_A(x)f_B(x)$
64. 假定 f 是从 A 到 B 的函数，A 和 B 为有限集，且 $|A|=|B|$。证明 f 是一对一的当且仅当它是映上的。
65. 证明或否定下列关于上取整函数和下取整函数的语句。
   a) 对任意实数 x，$\lceil\lfloor x\rfloor\rceil=\lfloor x\rfloor$。
   b) 只要 x 是实数，$\lfloor 2x\rfloor=2\lfloor x\rfloor$。
   c) 只要 x 和 y 是实数，$\lceil x\rceil+\lceil y\rceil-\lceil x+y\rceil=0$ 或 1。
   d) 对任意实数 x 和 y，$\lceil xy\rceil=\lceil x\rceil\lceil y\rceil$。
   e) 对任意实数 x，$\left\lceil\frac{x}{2}\right\rceil=\left\lfloor\frac{x+1}{2}\right\rfloor$。
66. 证明或否定下列关于下取整函数和上取整函数的语句。
   a) 对任意实数 x，$\lfloor\lceil x\rceil\rfloor=\lceil x\rceil$。　　b) 对任意实数 x 和 y，$\lfloor x+y\rfloor=\lfloor x\rfloor+\lfloor y\rfloor$。
   c) 对任意实数 x，$\lceil\lceil x/2\rceil\rceil=\lceil x/4\rceil$。　　d) 对任意实数 x，$\lfloor\sqrt{\lceil x\rceil}\rfloor=\lfloor\sqrt{x}\rfloor$。
   e) 对任意实数 x 和 y，$\lfloor x\rfloor+\lfloor y\rfloor+\lfloor x+y\rfloor\leqslant\lfloor 2x\rfloor+\lfloor 2y\rfloor$。
67. 证明若 x 是正实数，那么
   a) $\lfloor\sqrt{\lfloor x\rfloor}\rfloor=\lfloor\sqrt{x}\rfloor$　　b) $\lceil\sqrt{\lceil x\rceil}\rceil=\lceil\sqrt{x}\rceil$
68. 令 x 为实数，证明 $\lfloor 3x\rfloor=\lfloor x\rfloor+\left\lfloor x+\frac{1}{3}\right\rfloor+\left\lfloor x+\frac{2}{3}\right\rfloor$。

用来计算函数值的程序也许不能给出该函数在其所有元素上的正确值。例如程序可能因无限循环或内存溢出而不能给出某个正确值。为研究这种情况，我们用到*部分函数*的概念。从集合 A 到集合 B 的部分函数 f 是为 A 的一个子集的每个元素指定 B 中唯一值的指派。A 的这个子集称为 f 的定义区域。集合 A 和 B 分别称为 f 的*定义域*和*伴域*。我们说 f 在 A 中不属于定义区域的元素上未定义。我们用 $f: A\to B$ 表示 f 是从 A 到 B 的部分函数(这个符号与表示函数的符号一样。符号的上下文将决定 f 究竟是部分函数还是全函数)。当 f 的定义区域等于 A 时，就说 f 是*全函数*。
69. 对下列各个部分函数求它的定义域、伴域、定义区域及其未定义的值的集合。另外判断它是否为全

函数。

a) f：$\mathbf{Z}\to\mathbf{R}$，$f(n)=1/n$　　b) f：$\mathbf{Z}\to\mathbf{Z}$，$f(n)=\lceil n/2\rceil$

c) f：$\mathbf{Z}\times\mathbf{Z}\to\mathbf{Q}$，$f(m,n)=m/n$　　d) f：$\mathbf{Z}\times\mathbf{Z}\to\mathbf{Z}$，$f(m,n)=mn$

e) f：$\mathbf{Z}\times\mathbf{Z}\to\mathbf{Z}$，$f(m,n)=m-n$，如果 $m>n$

70. a)证明从 A 到 B 的部分函数 f 可以看成从 A 到 $B\cup\{u\}$ 的函数 f^*，其中 u 不是 B 的元素，且

$$f^*(a)=\begin{cases}f(a),\text{若 } a \text{ 属于 } f \text{ 的定义区域}\\ u,\text{若 } f \text{ 在 } a \text{ 未定义}\end{cases}$$

b)使用 a)中的构造法，找出练习 59 中各部分函数对应的 f^*。

*71. 证明函数 f：$\mathbf{Z}^+\times\mathbf{Z}^+\to\mathbf{Z}^+$ 是一对一且是映上的，其中 $f(m,n)=(m+n-2)(m+n-1)/2$。

关键术语和结果

逻辑(1.1～1.4节)

术语

命题：一个或成真或为假的语句。

真值：真或假。

$\neg p$(p 的否定)：与 p 的真值相反的命题。

逻辑运算符：用于组合命题的运算符。

复合命题：用逻辑运算符组合命题构造出的命题。

真值表：显示命题真值的表。

$p\vee q$(p 和 q 的析取)：除非 p 和 q 均为假，否则为真的命题。

$p\wedge q$(p 和 q 的合取)：只有在 p 和 q 均为真时才为真的命题。

$p\oplus q$(p 和 q 的异或)：当 p 和 q 中恰有一个为真时才为真的命题。

$p\to q$(p 蕴含 q)：只有在 p 为真而 q 为假时才为假的命题。

$p\to q$ 的逆蕴含：蕴含关系 $q\to p$。

$p\to q$ 的倒置蕴含：蕴含关系 $\neg q\to\neg p$。

$p\to q$ 的反蕴含：蕴含关系 $\neg p\to\neg q$。

$p\leftrightarrow q$(双蕴含)：只有 p 和 q 真值相同时才为真的命题。

字位：0 或 1。

布尔变量：以 0 或 1 为值的变量。

字位运算：一个或多个字位上的运算。

位串：一串字位。

按位运算：位串上的运算，对一个位串的字位和另一位串的对应字位作运算。

永真式：永远为真的复合命题。

矛盾：永远为假的复合命题。

可能式：有时成真有时为假的复合命题。

一致复合命题：为变量赋真值，使所有这些命题为真的复合命题。

逻辑等价：复合命题是逻辑等价的，如果它们总有同样的真值。

命题函数：谓词和变量的组合。

论域：命题函数中变量的定义域。

$\exists xP(x)$($P(x)$的存在量化)：当且仅当在论域中存在一个 x 使 $P(x)$为真时才为真的命题。

$\forall xP(x)$($P(x)$的全称量化)：当且仅当论域中的所有 x 值 $P(x)$均为真时才为真的命题。

自由变量：命题函数中未绑定的变量。

结果　1.2 节表 1-13、表 1-14 和表 1-15 给出逻辑等价关系。

证明方法（1.5 节）

术语

定理：可以证明为真的数学断言。

猜想：真值未知的数学断言。

证明：对定理为真的说明。

引理：用来证明其他定理的简单定理。

推论：可以作为刚刚证明的定理的后果而被证明的命题。

推理规则：用来从已知的断言得出结论的重言式蕴含式。

谬误：常常被不正确地用来得出结论而且是一种可能式的蕴含式。

循环论证或回避问题：其中一步或多步是基于被证明的命题的真值的推理。

空证明：基于 p 为假的事实而对蕴含式 $p \to q$ 的证明。

平凡证明：基于 q 为真的事实而对蕴含式 $p \to q$ 的证明。

直接证明：通过证明当 p 为真时 q 必然为真而进行的对蕴含式 $p \to q$ 的证明。

间接证明：通过证明当 q 为假时 p 必然为假而进行的对蕴含式 $p \to q$ 的证明。

归谬证明：基于蕴含式 $\neg p \to q$ 的真值（其中 q 是矛盾式）而对命题 p 为真的证明。

分情形证明：证明每个前件都蕴含结论因而对前件是命题的析取式的蕴含式的证明。

反例：使得 $P(x)$ 为假的元素 x。

构造性的存在性证明：证明具有特定性质的元素存在，通过显式的方式来寻找这样的元素。

非构造性的存在性证明：证明具有特定性质的元素存在，但不是显式地寻找这样的元素。

有理数：可以表示为两个整数 p 和 q（其中 $q \neq 0$）之比的数。

唯一性证明：证明具有特定性质的元素唯一地存在。

集合（1.6～1.7 节）

术语

集合：一组互不相同的对象。

公理：理论的一个基本假设。

悖论：逻辑上的不一致性。

集合的元素、成员：集合中的一个对象。

$\varnothing$（空集）：没有成员的集合。

全集：包含考虑中的所有对象的集合。

文氏图：一个或多个集合的一种图形表示。

$S=T$（集合相等）：S 和 T 有相同的元素。

$S \subseteq T$（S 是 T 的子集）：S 的每个元素也是 T 的元素。

$S \subset T$（S 是 T 的真子集）：S 是 T 的子集，且 $S \neq T$。

有限集：含 n 个元素的集合，其中 n 是非负整数。

无限集：不是有限集的集合。

$|S|$（S 的基数）：S 中元素的个数。

$P(S)$（S 的幂集合）：S 的所有子集的集合。

$A \cup B$（A 和 B 的并集）：包含那些至少属于 A 和 B 之一的元素的集合。

$A \cap B$（A 和 B 的交集）：包含那些既属于 A 又属于 B 的元素的集合。

$A-B$（A 和 B 的差集）：包含那些只属于 A 而不属于 B 的元素的集合。

$\overline{A}$（A 的补集）：全集中不属于 A 的元素的集合。

$A \oplus B$（A 和 B 的对称差）：包含恰属于 A 和 B 之一的那些元素的集合。

成员表：显示集合中元素的成员关系的表格。

结果 1.7 节表 1-21 给出集合恒等式。

函数（1.8 节）

术语

从 A 到 B 的函数：一种指派，为 A 中每个元素指派 B 中恰好一个元素。

f 的定义域：指集合 A，若 f 是从 A 到 B 的函数。

f 的伴域：指集合 B，若 f 是从 A 到 B 的函数。

b 是 f 之下 a 的像：$b=f(a)$。

a 是 f 之下 b 的原像：$f(a)=b$。

f 的值域：f 的像的集合。

映上函数，满射：从 A 到 B 的函数，使 B 的每个元素都是 A 中某元素的像。

一对一函数，内射：定义域中每个元素的像都不相同的函数。

一一对应，双射：既是一对一又是映上的函数。

f 的逆：（当 f 是双射时）颠倒由 f 给出的对应关系的函数。

$f \circ g$（f 和 g 的组合）：为 x 指派 $f(g(x))$ 的函数。

$\lfloor x \rfloor$（下取整函数）：不超过 x 的最大整数。

$\lceil x \rceil$（上取整函数）：大于或等于 x 的最小整数。

复习题

1. a）定义命题的否定。
   b）“这是一门烦人的课程”的否定是什么？
2. a）（用真值表）定义命题 p 和 q 的析取、合取、异或、蕴含和双蕴含。
   b）“今晚我去看电影”和“我将完成离散数学作业”的析取、合取、异或、蕴含和双蕴含是什么？
3. a）给出至少五种用汉语表达蕴含 $p \rightarrow q$ 的方式。
   b）定义蕴含的逆蕴含和倒置。
   c）“如果明天阳光明媚，我将到林中散步。”给出此蕴含语句的逆蕴含和倒置。
4. a）两个命题逻辑等价的含义是什么？
   b）描述证明两个复合命题逻辑等价的不同方法。
   c）至少用两种方法证明 $\neg p \vee (r \rightarrow \neg q)$ 和 $\neg p \vee \neg q \vee \neg r$ 等价。
5. （依赖于 1.2 节的练习）
   a）给出一个真值表，解释怎样用析取范式构造一个复合命题使其真值表就是你给出的真值表。
   b）解释为什么 a）说明运算符 $\wedge$，$\vee$ 和 $\neg$ 是功能完全的。
   c）是否有一个运算符使得只含这个运算符的集合是功能完全的？
6. 什么是谓词 $P(x)$ 的全称和存在量化？什么是它们的否定？
7. a）量化语句 $\exists x \forall y P(x, y)$ 和 $\forall y \exists x P(x, y)$ 的区别何在？其中 $P(x, y)$ 为谓词。
   b）给出谓词 $P(x, y)$ 的一个实例，使 $\exists x \forall y P(x, y)$ 和 $\forall y \exists x P(x, y)$ 有不同的真值。
8. a）描述一下蕴含式 $p \rightarrow q$ 的直接证明、间接证明和归谬证明分别是什么意思。
   b）分别给出命题：“若 n 是偶数，则 $n+4$ 是偶数”的直接证明、间接证明和归谬证明。
9. a）描述证明双向条件 $p \leftrightarrow q$ 的一种方式。
   b）证明命题：“整数 $3n+2$ 是奇数当且仅当整数 $9n+5$ 是偶数，其中 n 是整数”。
10. 为了证明命题 p_1，p_2，p_3 和 p_4 都是等价的，证明蕴含式 $p_4 \rightarrow p_2$，$p_3 \rightarrow p_1$ 和 $p_1 \rightarrow p_2$ 都是有效的，这

样做是否充分？若不充分，则给出可用来证明这四个命题都是等价的另外一组蕴含式。

11. a)假定形如 $\forall xP(x)$ 的蕴含式为假。如何可以证明它为假？
    b)证明命题“对每个正整数 n 来说，$n^2 \geqslant 2n$”为假。
12. a)构造性与非构造性存在性证明之间的差异是什么？对这两种证明分别举一个例子。
13. 证明存在唯一的元素 x 使得 $P(x)$ 为真(其中 $P(x)$ 是命题函数)。这种证明的基础是什么？
14. a)定义两个集合的并集、交集、差集和对称差。
    b)什么是正整数集合和奇整数集合的并集、交集、差集和对称差？
15. a)定义两个集合相等。
    b)描述证明两个集合相等的方法。
    c)至少给出证明 $A-(B\cap C)$ 和 $(A-B)\cup(A-C)$ 相等的两种不同方法。
16. 解释逻辑等价和集合恒等之间的关系。
17. a)定义集合 S 的基数 $|S|$。
    b)给出计算 $|A\cup B|$ 的一个公式，其中 A 和 B 为集合。
18. a)定义集合 S 的幂集合。
    b)集合 S 的幂集合何时为空集？　　c)含 n 个元素的集合 S 的幂集合有多少个元素？
19. a)定义函数的定义域、伴域和值域。
    b)令 $f(n)=n^2+1$ 为从整数集合到整数集合的函数。它的定义域、伴域和值域是什么？
20. a)给出从正整数集合到正整数集合的函数为一对一函数的含义。
    b)给出从正数集合到正整数集合的函数为映上函数的含义。
    c)给出一个从正整数集合到正整数集合的函数的实例，使它既是一对一的，又是映上的。
    d)给出一个从正整数集合到正整数集合的函数的实例，使它是一对一的，但不是映上的。
    e)给出一个从正整数集合到正整数集合的函数实例，使它不是一对一的，但是映上的。
    f)给出一个从正整数集合到正整数集合的函数实例，使它既不是一对一的，也不是映上的。
21. a)定义函数的反函数。
    b)什么样的函数有反函数？
    c)从整数集合到整数集合的函数 $f(n)=10-n$ 有反函数吗？如果有，其反函数是什么？
22. a)定义从实数集合到整数集合的下取整函数和上取整函数。
    b)对哪些实数 x 有 $\lfloor x\rfloor=\lceil x\rceil$?

补充练习

1. 令 p 为命题“我将完成本书中的每一道练习”，q 为命题“这门课程我将得‘A’”。将下列各项表示为 p 和 q 的组合。
    a)只有完成本书中的每道练习，这门课程我才能得‘A’。
    b)这门课程我将得‘A’，而且我将完成本书中每一道练习。
    c)或者这门课程我将得‘A’，或者我不完成本书中的每一道练习。
    d)这门课程我得‘A’的充分必要条件是我完成本书中每一道练习。
2. 求复合命题 $(p\vee q)\rightarrow(p\wedge\neg r)$ 的真值表。
3. 证明下列命题为永真式。
    a)$(\neg q\wedge(p\rightarrow q))\rightarrow\neg p$　　b)$((p\vee q)\wedge\neg p)\rightarrow q$
4. 给出下列蕴含关系的逆蕴含、反蕴含和倒置。
    a)如果今天下雨，我就开车上班。　b)如果 $|x|=x$，那么 $x\geqslant 0$。　c)若 n 大于 3，那么 n^2 大于 9。
5. 用命题变量 p，q，r 和 s 构造一个复合命题，使它在这些命题变量中恰有三个为真时取真值，其他情况下为假。

6. 证明下列语句是不一致的："如果 Sergei 拿到工作签证，那他将获得签约奖金。""如果 Sergei 拿到工作签证，那他将拿到高薪。""如果 Sergei 拿到签约奖金，那他将拿不到高薪"。"Sergei 拿到了工作签证。"
7. 证明下列语句是不一致的："如果 Miranda 没有修过离散数学课程，那她将不能毕业。""如果 Miranda 不能毕业，那她将不能胜任那份工作。""如果 Miranda 读了这本书，那她将能胜任那份工作。""Miranda 没有修过离散数学课程，但她读过这本书。"
8. 假定你在 1.1 节例 15 中描述的"武士和流氓"共居的岛上遇见三个人 A、B 和 C。如果 A 说"我是流氓，B 是骑士"且 B 说"我们三人中恰有一人是骑士"，那么 A、B 和 C 分别是哪种人？
9. (摘自[Sm78])假定在一个岛上住着三类人：武士、流氓和普通人。武士总是说真话，流氓总是说谎话，普通人有时说谎话有时说真话。侦探为了找出罪犯，询问了岛上的三个人 Amy、Brenda 和 Claire。侦探知道三人中有一人犯罪了，但不知是哪个人。他们也知道罪犯是一个骑士，另两个人不是骑士。此外，侦探还记录了如下供述。Amy 说："我是清白的。"Brenda 说："Amy 说的是真的。"Claire 说："Brenda 不是普通人。"经过分析，侦探找到了罪犯。他是谁？
10. 令 $P(x)$ 为语句"学生 x 会微积分"，$Q(y)$ 为"y 班上有个学生会微积分"。用 $P(x)$ 和 $Q(y)$ 的量化表示下列各项。
    a)某个学生会微积分。
    b)不是每个学生都会微积分。
    c)每个班上都有一个学生会微积分。
    d)每个班上的每个学生都会微积分。
    e)至少有一个班没有学生会微积分。
11. 令 $P(m, n)$ 为语句"m 除尽 n"，其中变量 m 和 n 的论域均为正整数集合。给出下列命题的真值。
    a) $P(4, 5)$
    b) $P(2, 4)$
    c) $\forall m \forall n P(m, n)$
    d) $\exists m \forall n P(m, n)$
    e) $\exists n \forall m P(m, n)$
    f) $\forall n P(1, n)$
12. 用量词表达语句"没人有多于三个的祖母"，使用命题函数 $G(x, y)$，它表示"x 是 y 的祖母。"
13. 用量词表达语句"每个人恰有两个亲生父母"，使用命题函数 $P(x, y)$，它表示"x 是 y 的亲生父母。"
14. 令 $P(x, y)$ 为命题函数。证明蕴含关系 $\exists x \forall y P(x, y) \rightarrow \forall y \exists x P(x, y)$ 为永真式。
15. 令 $P(x)$ 和 $Q(x)$ 为命题函数。求证 $\exists x(P(x) \rightarrow Q(x))$ 和 $\forall x P(x) \rightarrow \exists x Q(x)$ 有同样的真值。
16. 如果 $\forall y \exists x P(x, y)$ 为真，$\exists x \forall y P(x, y)$ 是否必定也为真？
17. 如果 $\forall x \exists y P(x, y)$ 为真，$\exists x \forall y P(x, y)$ 是否必定也为真？
18. 找出下列语句的否定。
    a)如果今天下雪，那么我明天去滑雪。
    b)班上每个人都懂数学归纳法。
    c)班上有些学生不喜欢离散数学。
    d)每堂数学课都有某个学生上着课就睡着了。
19. 用量词表示"班上每个学生都选修过数学学院每个系的某门课"。
20. 用量词表示"在美国某学院的校园里有座楼的每间屋子都漆成了白色。"
21. 令 A 为英语中含字母 x 的所有单词的集合，B 是英语中含字母 q 的所有单词的集合。把下列各集合表示为 A 和 B 的组合。
    a)英语中不含字母 x 的单词的集合。
    b)英语中既含字母 x 又含字母 q 的单词的集合。
    c)英语中含字母 x 但不含字母 q 的单词的集合。
    d)英语中不含字母 x 或字母 q 的单词的集合。
    e)英语中含 x 或含 q 但不同时含 x 和 q 的单词的集合。
22. 描述一个推理规则，可用它来证明论域中恰有两个元素 x 和 y 使得 $P(x)$ 和 $P(y)$ 为真。用汉语句子表达这个推理规则。
23. 下述论证错在哪里？给定前提 $\exists x P(x) \wedge \exists x Q(x)$，使用化简得到 $\exists x P(x)$；使用存在例示得到 $P(c)$（对某些元素 c）；再使用化简得到 $\exists x Q(x)$；使用存在例示得到 $Q(c)$（对某些元素 c）；使用合取得出 $P(c) \wedge Q(c)$；最后，使用存在生成得出 $\exists x(P(x) \wedge Q(x))$。
24. 使用推理规则证明前提 $\forall x(P(x) \rightarrow Q(x))$、$\forall x(Q(x) \rightarrow R(x))$ 和 $\neg R(a)$ 蕴含结论 $\neg P(a)$，其中 a 在论域中。
25. 使用推理规则证明前提 $\forall x(P(x) \wedge Q(x))$ 和 $\forall x(\neg P(x) \wedge Q(x) \rightarrow R(x))$ 蕴含结论 $\forall x(\neg R(x) \rightarrow P(x))$。
26. 证明若 x^3 是无理数，则 x 是无理数。

27. 证明若 x 是无理数且 $x\geqslant 0$，则$\sqrt{x}$是无理数。
28. 证明给定一个非负整数 n，存在一个唯一的非负整数 m 使得 $m^2\leqslant n<(m+1)^2$。
29. 证明存在一个整数 m 使得 $m^2>10^{1000}$。你的证明是构造性的还是非构造性的？
30. 证明存在这样的正整数：它可以被以两种不同的方式写成正整数的立方和。使用计算机或计算器来完成。
31. 反驳如下命题：每个正整数均可表示为 8 个非负整数的立方和。
32. 反驳如下命题：每个正整数均可表示为至多两个非负整数的平方与一个非负整数的立方的和。
33. 反驳如下命题：每个正整数均可表示为 36 个非负整数的 5 次方的和。
34. 求证：若 A 是 B 的子集，则 A 的幂集是 B 的幂集的子集。
35. 假设集合 A 的幂集是集合 B 的幂集的子集。A 是否一定是 B 的子集？
36. 令 **E** 表示偶数集合，**O** 表示奇数集合，**Z** 依旧表示所有整数的集合。求下列各项：
    a) $\mathbf{E}\cup\mathbf{O}$　　b) $\mathbf{E}\cap\mathbf{O}$　　c) $\mathbf{Z}-\mathbf{E}$　　d) $\mathbf{Z}-\mathbf{O}$
37. 如果 A 是集合，U 是全集，求证
    a) $A\cap\overline{A}=\varnothing$　　b) $A\cup\overline{A}=U$
38. 若 A 和 B 为集合，求证
    a) $A=A\cap(A\cup B)$　　b) $A=A\cup(A\cap B)$
39. 求证：若 A，B 为集合，则 $A-(A-B)=A\cap B$。
40. 令 A，B 为集合。求证 $A\subseteq B$ 当且仅当 $A\cap B=A$。
41. 令 A，B，C 为集合，求证$(A-B)-C$不一定等于$A-(B-C)$。
42. 假定 A，B，C 为集合。证明或否定$(A-B)-C=(A-C)-B$。
43. 假定 A，B，C，D 为集合。证明或否定$(A-B)-(C-D)=(A-C)-(B-D)$。
44. 证明若 A，B 为有限集合，则 $|A\cap B|\leqslant|A\cup B|$。什么条件下这一关系成为等式？
45. 令 A，B 为有限全集 U 中的集合。按从小到大的顺序排列下面各组数。
    a) $|A|$，$|A\cup B|$，$|A\cap B|$，$|U|$，$|\varnothing|$
    b) $|A-B|$，$|A\oplus B|$，$|A|+|B|$，$|A\cup B|$，$|\varnothing|$
46. 令 A，B 为有限全集 U 的集合。求证 $|\overline{A}\cap\overline{B}|=|U|-|A|-|B|+|A\cap B|$。
47. 令 f 和 g 分别是从$\{1, 2, 3, 4\}$到$\{a, b, c, d\}$和从$\{a, b, c, d\}$到$\{1, 2, 3, 4\}$的函数，使得 $f(1)=d$，$f(2)=c$，$f(3)=a$，$f(4)=b$ 和 $g(a)=2$，$g(b)=1$，$g(c)=3$，$g(d)=2$。
    a) f 是一对一的吗？g 是一对一的吗？　　b) f 是映上的吗？g 是映上的吗？
    c) f 或 g 是否有反函数？若有，求其反函数。
48. 令 f 为从集合 A 到集合 B 的一对一函数。令 S 和 T 为 A 的子集。求证 $f(S\cap T)=f(S)\cap f(T)$。
49. 给出一个例子说明如果 f 不是一对一的，那么练习 48 中的等式可能不成立。
50. 求证：若 n 为整数，则 $n=\lceil n/2\rceil+\lfloor n/2\rfloor$。
51. 对实数 x 和 y，什么情况下 $\lfloor x+y\rfloor=\lfloor x\rfloor+\lfloor y\rfloor$为真？
52. 对实数 x 和 y，什么情况下$\lceil x+y\rceil=\lceil x\rceil+\lceil y\rceil$为真？

计算机题目

按给定的输入和输出写程序。

1. 已知命题 p 和 q 的真值，求它们的合取、析取、异或、蕴含和双蕴含的真值。
2. 已知长度为 n 的两个位串，求它们的按位 AND、按位 OR 及按位 XOR。
3. 已知模糊逻辑中命题 p 和 q 的真值，求 p 和 q 的析取和合取的真值(参看 1.1 节练习 35～37)。
4. 已知含 n 个元素的某集合的子集 A 和 B，用位串求 $\overline{A}$，$A\cup B$，$A\cap B$，$A-B$ 和 $A\oplus B$。
5. 已知同一全集的多重集 A 和 B，求 $A\cup B$、$A\cap B$、$A-B$ 和 $A+B$(参看 1.7 节练习 49 的开场白)。
6. 已知模糊集合 A 和 B，求 $\overline{A}$，$A\cup B$ 和 $A\cap B$(参看 1.7 节练习 51 的开场白)。

7. 已知从{1，2，…，n}到整数集合的函数 f，判断 f 是否是一对一的。
8. 已知从{1，2，…，n}到它自己的函数 f，判断 f 是否是映上的。
9. 已知从{1，2，…，n}到它自己的双射 f，求 f^{-1}。

计算和研究

使用一个计算程序或你已完成的程序做下面的练习。

1. 找出这样的正整数：它不是 9 个不同的正整数的立方和。
2. 找出大于 79 的这样的正整数：它不是 18 个正整数的四次方的和。
3. 找出尽可能多这样的正整数：它可以用正整数的立方和来表示。用两种不同的方式，对 1729 使用这个性质。
4. 计算从集合 S 到集合 T 的一对一函数的个数，其中 S 和 T 为各种大小的有限集合。能给出计算这种函数个数的公式吗？（在第 4 章有这样一个公式。）
5. 计算从集合 S 到集合 T 的映上函数的个数，其中 S 和 T 是各种大小的有限集合。能给出计算这种函数个数的公式吗？（第 6 章有这样一个公式。）

写作题目

用课本以外的资料，按下列要求写成短文。

1. 讨论如下逻辑悖论：克里特人 Epimenides 的悖论（又叫说谎者悖论）、Jourdain 的纸牌悖论以及理发师悖论。说明如何解决它们。
2. 描述模糊逻辑怎样用于实际应用。参考为一般读者写的一本或几本最近出版的模糊逻辑书。
3. 描述由 Layman Allen 提出的“WFF'N PROOF，The Game of Modern Logic”的基本规则。给出 WFF'N PROOF 中包含的一些博弈示例。
4. 读一些卡罗尔(Lewis Carroll)关于符号逻辑的作品。详细描述他用于表示逻辑论证和推理规则的一些模型。
5. 扩展 1.3 节对 Prolog 的讨论，进一步解释 Prolog 如何使用消解。
6. 讨论在计算逻辑中使用的一些技术，包括 Skolem 的规则。
7. “自动定理证明”是使用计算机来机械地证明定理的任务。讨论自动定理证明的目标和应用，以及在开发自动定理证明器上取得的进步。
8. 讨论如何建立一个公理集合论才能避免罗素悖论。（参看 1.6 节练习 30。）
9. 研究函数概念是从哪里首先提出来的，描述一下这一概念最初是怎样应用的。



第2章　基础：算法、整数和矩阵

许多问题都可以当作一般问题的特例来解决。例如，寻找序列 101，12，144，212，98 中最大整数的问题。这是寻找整数序列中最大整数问题的一个特例。为解决这个一般问题，必须给出一个算法。这个算法规定解决一般问题的一系列步骤。本书将介绍解决许多不同类型问题的算法，例如，求两个整数最大公约数的算法、产生有限集所有排序的算法、搜索列表的算法、搜索序列元素的算法、求网络上两顶点间最短路径的算法等。

与算法有关的一项重要考虑是算法的计算复杂度，即用这一算法解决一定规模的问题需要什么计算机资源。为了度量算法的复杂度，使用大 O 和大 Θ 记号。本章将介绍这些记号，对算法的复杂度进行分析，另外将讨论就实践和理论而言算法的复杂度意味着什么。

整数集合是离散数学的基础。特别是整数除法的概念是计算机算术的基础。我们将简要回顾一下数论(整数及其性质的研究)的几个重要概念。学习与整数有关的几个重要算法，包括求最大公约数的欧几里得算法，这是几千年前就给出的算法。整数可以用任何大于 1 的正整数为基数来表示。贯穿计算机科学的二进制展开即是以 2 为基数的表示。本章将学习以 b 为基数的整数表示，并给出算法求这些表示。还要讨论用于整数算术的算法，这是第一批被称为算法的过程。本章还要介绍几个重要的数论应用。例如，本章将利用数论来进行信息保密、生成伪随机数并为计算机文件分配内存地址。一度被认为是最纯粹数学科目的数论已成为计算机和网络安全的实质性工具。

矩阵在离散数学中用于表示各种离散结构。我们将学习矩阵的基本内容及表示关系和图形所需的矩阵算术。矩阵算术将用于与这些结构有关的大量算法。

2.1　算法

2.1.1　引言

离散数学中有多种一般性的问题。例如，已知一串整数，求最大的一个；已知一个集合，列出其所有子集；给定一个整数集合，把这些整数从小到大排序；已知一个网络，找两个顶点之间的最短路径等。遇到这样的问题时，首先要做的就是构造一个模型，把问题翻译为数学问题。在这种模型中用到的离散结构包括：集合、序列和函数。这些是第 1 章讨论过的结构，以及置换、关系、图、树、网络和有穷状态机等其他结构。这些概念将在以后各章讨论。

建立合适的数学模型只是解题的第一步。完整的解题还需要利用这一模型解决一般性问题的方法。理想的情况是求一个过程。这个过程能够遵循一串步骤找出所求的答案。这一串步骤称为算法。

定义 1　算法是进行一项计算或解决一个问题的一组有限多条准确的指令。

算法(algorithm)这一术语是对 al-khowarizmi(花拉子密)[⊖]这一名字的讹用。花拉子密是 9 世

⊖ **Abu Ja'far Mohammed ibn Musa al-khowarizmi(约公元 780—850)**　天文学家和数学家花拉子密是巴格达一个科学家组织“智慧之家”的成员。花拉子密这一名字的含义是“来自花拉子模(Kowarzizm)镇”。该镇现称卡瓦(Khiva)，是乌兹别克斯坦的一部分。花拉子密写过关于数学、天文学和几何学的书。西欧人首先从他的著作中学习代数。代数(algebra)一词源自 al-jabr，这是他的书的标题《Kitab al-jabr w'al muquabala》的一部分。这本书曾译为拉丁文并广泛用作课本。他关于印度数字的书描述了使用这些数字作算术运算的过程。欧洲作者使用了他的名字的一个拉丁讹音来表示用印度数字做算术的内容，后来演变为 algorithm(算法)一词。

纪的一位数学家。他关于印度数字的书是现代十进制记号的基础。起初 algorism 这个词用于表示使用十进制记号做算术的规则。18 世纪时 algorism 进化为 algorithm。随着人们对计算机的兴趣的增长，算法的概念有了更广的含义，已经不仅包含做算术的过程，而且包含所有确定的解题过程。（在 2.5 节将讨论做整数算术的算法。）

本书将讨论解决各种问题的算法。这一节用求有限整数序列中最大整数这一问题解释算法的概念和算法具有的性质。还要给出在有限集合中找一个特定元素的算法。以后各节将讨论求两个整数的最大公约数的过程，求网络上两点之间的最短路径的过程，矩阵相乘的过程等。

例 1 求有限整数序列中最大值的算法。

Extra Examples 尽管求有限序列最大元素的问题相对简单，但这个问题能很好地说明算法概念。另外，需要求有限整数序列中最大整数的情况各不相同。例如，大学可能要找出几千名学生参加竞赛的最高分。体育组织可能要确定每月成绩最好的运动员。人们希望开发出一个算法。求有限整数序列最大元素的问题就可以使用该算法。

可以用几种不同的方式给出解决这一问题的过程。一种方法是直接用中文描述使用的一串步骤。现在就给出这样一个解。

例 1 的解：采取下面的步骤。

1. 设临时最大值等于序列中第一个整数。（在本过程的每一阶段，临时最大值都等于已检查过的最大整数。）

2. 将序列中的下一整数与临时最大值比较，如果这个数大于临时最大值，置临时最大值为这一整数。

3. 如果序列中还有其他整数，重复前一步骤。

4. 在序列中没有留下可比的整数时停止，此刻的临时最大值就是序列中的最大整数。 ■

也可以用计算机语言来描述算法。但是当这样做时，只能使用这种语言所允许的指令。这样做常常导致复杂而难以理解的算法描述。另外，由于许多程序设计语言都是常用的，并不希望从中选用特定的某种语言。因此本书不用任何一种特定的计算机语言描述算法，而是使用伪码的形式。（本书中所有算法也都用英文描述。）伪码提供的是在算法的英文描述及其程序设计语言实现之间的中间一步。算法步骤用模仿程序设计语言指令的伪指令描述，不过伪指令可以包括有确切定义的一切运算和语句。以算法的伪码描述为起点，可以用任何一种计算机语言产生计算机程序。

本书使用的伪码大致上以程序设计语言 Pascal 为基础。不过并不遵循 Pascal 或其他程序设计语言的语法。此外，任何有确切定义的指令均可用于伪码中，附录 B 给出了本书使用的伪码的细节。必要时读者应参看这一附录。

下面是求有限序列最大元素的伪码算法。

算法 1 求有限序列的最大元素

```
procedure max(a1, a2, …, an: 整数)
max := a1
for i := 2 to n
    if max < ai then max := ai
{max 是最大元素}
```

这一算法首先把序列的首项 a_1 赋给变量 max。“for”循环用于逐个检查序列项。如果某一项大于 max 的当前值，就将其赋给 max，成为 max 的新值。

算法有几个通用的性质。在描述算法时记住这些性质是有用的。这些性质是：

- **输入** 算法从一个指定的集合得到输入值。
- **输出** 对每个输入值集合，算法都要从一个指定的集合中给出输出值。输出值就是问题的解。
- **确定性** 算法的步骤必须是准确定义的。
- **正确性** 对每一组输入值，算法都应产生正确的输出值。
- **有限性** 对任何输入算法都应在有限(可能很多)步之后产生所求的输出。
- **有效性** 算法的每一步都应能够准确地在有限时间内完成。
- **通用性** 算法过程应该可以应用于期望形式的所有问题，而不只是用于特定的输入值。

例 2 证明求有限整数序列最大元素的算法 1 具有上面列出的所有性质。

解 算法 1 的输入是个整数序列。输出是该序列的最大整数。算法的每一步都是准确定义的，因为只出现赋值、有限循环和条件语句。为了说明算法是正确的，必须说明当算法终止时，变量 max 的值等于序列的最大项。为了看明白这一点，注意 max 的初值是序列的第一项；当检查序列的后继各项时，如果有一项超过已检查项的最大值，就把 max 更新为这一项的值。这个(非形式化)论证说明当检查了所有的项时，max 就等于最大项的值。(这个事实的严格证明需要 3.3 节介绍的技术。)算法使用有限的步骤数，因为在检查过该序列中所有整数以后算法就终止。算法在有限时间内完成，因为每一步要么是比较，要么是赋值。最后，算法 1 是通用的，因为该算法可用于求任何有限整数序列的最大元素。 ■

2.1.2 搜索算法

在有序表中为元素定位的问题常会出现。例如检查单词拼写的程序要在字典中搜索，字典其实就是单词的有序表。这一类问题称为*搜索问题*。本节将讨论几个搜索算法。在 2.2 节将介绍这些算法中各自使用的步骤数。

一般的搜索问题可以描述如下：在不同元素 a_1，a_2，…，a_n 的表中为元素 x 定位，或判定 x 不在该表中。这一搜索问题的解是表中等于 x 的那一项的位置(即，若 $x=a_i$，那么 i 就是解)，或当 x 不在表中时解为 0。

线性搜索 将介绍的第一个算法称为*线性搜索*(或*顺序搜索*)算法。线性搜索算法从比较 x 和 a_1 开始。若 $x=a_1$，那么解就是 a_1 的位置，也就是 1。当 $x\neq a_1$ 时，比较 x 和 a_2。若 $x=a_2$，解就是 a_2 的位置，也就是 2。当 $x\neq a_2$ 时，比较 x 与 a_3。继续这一过程，逐一比较 x 和表中的每一项；除非不出现相等，否则一旦发现相等，解就是该项的位置。如果已搜索了整个表却不能为 x 定位，那么解是 0。这一线性搜索算法的伪码如算法 2 所示。

算法 2 线性搜索算法

```
procedure linear search(x: 整数, a1, a2, …, an: 不同整数)
i := 1
while(i ≤ n 和 x ≠ ai)
    i := i+1
if i ≤ n then location := i
else location := 0
{location 是等于 x 的项的下标，或是 0(若找不到 x)}
```

二分搜索 现在考虑另一个搜索算法。当表中各项以升序出现时可以用这一算法。(例如，若各项为数，则按从最小到最大顺序排列；如果各项是单词，则可以按字典序或字母序排列。)这

第二个算法称为二分搜索算法。这个算法比较要搜索的元素与表的中间项。然后此表就分成两个较小的长度相等的子表，或许其中有一个比另一个少一项。根据与中间项的比较结果，搜索将限于在两个子表中的一个进行。在 2.3 节将证明二分搜索算法比线性搜索算法的效率高得多。下面的例子说明如何进行二分搜索。

例 3 要在表

1,2,3,5,6,7,8,10,12,13,15,16,18,19,20,22

中搜索 19，第一步把有 16 个项的这个表分成各含 8 项的两个较小的表，即

1,2,3,5,6,7,8,10　　12,13,15,16,18,19,20,22

然后比较 19 和头一个表的最大项。因为 $10<19$，对 19 的搜索可以限于包含原表第 9～16 项的表中。下一步把含 8 项的这个表分成两个含 4 项的小表，即

12,13,15,16　　18,19,20,22

因为 $16<19$(将 19 与第一个表的最大项比较)，搜索限于这两个表的第二个，它包含原表的第 13～16 项。表 18，19，20，22 再分成两个，即

18,19　　20,22

因为 19 不大于两个表中第一个的最大项，这最大项也是 19，搜索限于第一个表：18，19。这个表包含原表的第 13、14 项。下一步，这个两项的表分成各含一项的两个表 18 和 19。因为 $18<19$，搜索限于第二个表：这是只含原表第 14 项，即 19 的表。现在搜索已经压缩到一项上。与这一项比较，19 定位为原表的第 14 项。■

现在给出二分搜索算法的步骤。要在表 a_1，a_2，…，a_n 中搜索整数 x，其中 $a_1<a_2<\cdots<a_n$，从比较 x 和序列的中间项 a_m 开始，其中 $m=\lfloor(n+1)/2\rfloor$。(回忆一下，$\lfloor x\rfloor$是不超过 x 的最大整数。)如果 $x>a_m$，搜索可以限制在序列的下半段，即 a_{m+1}，a_{m+2}，…，a_n。如果 x 不大于 a_m，搜索可限制在序列的上半段，即 a_1，a_2，…，a_m。

现在搜索的范围限于一个不超过$\lceil n/2\rceil$个元素的表。用同样的过程，比较 x 和这个短表的中间项，然后把搜索限于短表的前半段或后半段，这样重复直到得到只含一项的表。然后判断这项是否就是 x。二分搜索算法的伪码以算法 3 给出。

算法 3 二分搜索算法

```
procedure binary search(x: 整数, a1, a2, …, an: 递增整数)
i := 1  {i 是搜索区间的左端点}
j := n  {j 是搜索区间的右端点}
while i<j
begin
    m := ⌊(i+j)/2⌋
    if x>am then i := m+1
    else j := m
end
if x=ai then location := i
else location := 0
{location 是等于 x 的项的下标，或是 0(若找不到 x)}
```

算法 3 一次次压缩被搜索的序列。在任何阶段都只有从 a_i 开始到 a_j 结束的这些项需要考

虑。换言之，i 和 j 分别是留下需检查的最小和最大下标。算法 3 不断压缩需搜索的序列，直到剩下只有一项的序列。当做到这一步时，再比较看这一项是否等于 x。

2.1.3 排序

对一个表中的元素排序是一个常见问题。例如，为了制作电话簿，就要按字母顺序排列用户姓名。在电子邮件列表中给地址排序，可确定有无重复地址。建立有用的字典，需要把单词按字母顺序排列。同样，生成一份零件表，需要按照从小到大的零件号排列零件。

假定有一个集合的元素表。还假设有一种方式给集合的元素排序。(给集合元素排序的概念将在 7.6 节详细讨论。)排序就是把这些元素排成一个表，表中元素按照升序排列。例如，对表 7，2，1，4，5，9 的排序就产生表 1，2，4，5，7，9。对表 d，h，c，a，f 的排序(利用字母序)就产生表 a，c，d，f，h。

由于大量的计算资源用于对各种各样的东西排序，因此人们投入了很多努力来开发排序算法。用不同的策略设计数目多得惊人的排序算法。在《计算机程序设计艺术》(*The Art of Computer Programming*)这部重要著作中，高德纳(Donald Knuth)用了近 400 页篇幅叙述排序，详细讨论了大约 15 种不同的排序算法！计算机科学家和数学家对排序算法感兴趣的原因有：一些算法更容易实现，一些算法更有效(在一般情况下，或在已知输入带有某种特征时，比如稍有错位的表)，一些算法利用特殊的计算机体系结构，一些算法特别智能等。本节介绍两种排序算法：冒泡排序和插入排序。本节末尾的练习中介绍另外两种排序算法：选择排序和二分插入排序。本章末尾的"补充练习"中介绍剃须刀(shaker)排序。在 3.5 节将讨论归并排序并在该节末尾的练习中介绍快速排序；在 9.2 节练习中介绍竞赛排序。本书讨论排序算法，是因为排序是一个重要问题，还因为这些算法可作为许多重要概念的例子。

冒泡排序 冒泡排序是最简单的排序算法之一，但不是最有效的排序算法之一。冒泡排序把一个表排列成升序：一次次比较相邻的元素，若相邻元素顺序不对，就交换相邻元素。为了完成冒泡排序，执行基本操作，即交换一个较大元素与其后的较小元素，从表的头部开始，对整个表执行一遍。让这个过程迭代，直到排序宣告完成为止。算法 4 给出冒泡排序的伪码。可以想象把表中的元素排成一列。在冒泡排序里，当较小的元素与较大的元素交换时，较小的元素就"冒泡"到顶上。较大的元素则"下沉"到底下。例 4 对此进行说明。

例 4 用冒泡排序把 3，2，4，1，5 排列成升序。

解 图 2-1 说明这个算法的步骤。首先比较前两个元素 3 和 2。因为 3>2，交换 3 与 2，产

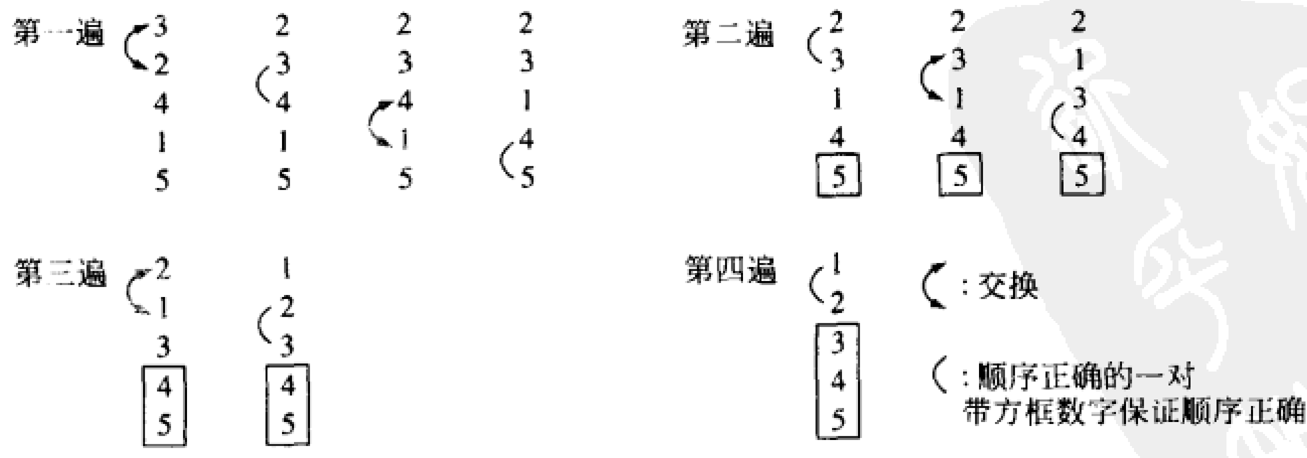


图 2-1 冒泡排序的步骤

生表2，3，4，1，5。因为3<4，继续比较4和1。因为4>1，交换4与1，产生表2，3，1，4，5。因为4<5，第一遍就完成了。第一遍保证了最大元素5是在正确位置上。

第二遍首先比较2和3。因为这两个数顺序正确，就比较3和1。因为3>1，交换这两个数，产生2，1，3，4，5。因为3<4，这两个数顺序正确。这一遍不必进一步比较了，因为5已经在正确位置上。第二遍保证两个最大元素4和5都在正确位置上。

第三遍首先比较2和1。因为2>1，交换这两个数，产生1，2，3，4，5。因为2<3，所以这两个数顺序正确。这一遍不必进一步比较了，因为4和5都已经在正确位置上。第三遍保证三个最大元素3，4和5都在正确位置上。

第四遍包括一次比较，即1和2的比较。因为1<2，所以这两个数顺序正确。这样就完成了冒泡排序。■

算法 4 冒泡排序

procedure *bubblesort*(a_1，…，a_n)
for $i:=1$ **to** $n-1$
　for $j:=1$ **to** $n-i$
　　if $a_j>a_{j+1}$ **then** 交换 a_j 与 a_{j+1}
{a_1，…，a_n 按升序排列}

插入排序 插入排序是一种简单的排序算法，但通常不是最有效的。为了给 n 个元素的表排序，插入排序从第二个元素开始。插入排序比较第二个元素与第一个元素。若第二个元素不超过第一个元素，把第二个元素插入到第一个元素前面；若第二个元素超过第一个元素，把第二个元素插入到第一个元素后面。此时前两个元素顺序正确。然后第三个元素与第一个元素比较，若第三个元素大于第一个元素，第三个元素与第二个元素比较；第三个元素插入到前三个元素中的正确位置上。

一般来说，在插入排序第 j 步上，表的第 j 个元素插入到已经排序的 $j-1$ 个元素的表的正确位置上。为了在表中插入第 j 个元素，使用线性搜索技术(参见练习43)；从表的头部开始，第 j 个元素与已经排序的 $j-1$ 个元素依次比较，直到发现第一个不小于这个元素的元素为止，或者直到这个元素已经与所有 $j-1$ 个元素都比较过为止；第 j 个元素插入到正确位置上，使得前 j 个元素排好顺序。这个算法继续下去，直到最后一个元素相对于前 $n-1$ 个元素已经排序的表插入到正确位置上。算法5用伪码描述插入排序。

例 5 用插入排序把表3，2，4，1，5排列成升序。

解 插入排序首先比较2和3。因为3>2，把2插入前一个位置，产生表2，3，4，1，5。接着比较4>2和4>3，把第三个元素4插入表的已排序部分。因为4>3，把4插入第三个位置上。此时表是2，3，4，1，5，我们知道前三个元素的顺序是正确的。接下来求第四个元素1在已排序元素2，3，4中的正确位置。因为1<2，得到表1，2，3，4，5。最后依次比较5与1，2，3和4，把5插入正确位置。因为5>4，5到了表的尾部，于是整个表都排成了正确的顺序。■

```
算法 5  插入排序
procedure insertion sort(a_1, a_2, …, a_n: 实数, n≥2)
for j := 2 to n
begin
    i := 1
    while a_j > a_i
        i := i+1
    m := a_j
    for k := 0 to j-i-1
        a_{j-k} := a_{j-k-1}
    a_i := m
end{a_1, a_2, …, a_n 已排序}
```

2.1.4 贪心算法

本书将学习的许多算法都设计用于解决优化问题。这种问题的目标是求给定问题把某个参数值最小化或最大化的解。本书后面学习的优化问题包括：求两个城市之间总长度最短的路线，确定一种方式用尽可能少的位数编码信息，以及求网络结点之间使用最少量光纤的一组光纤连接。

奇怪的是，一种最简单的方法常常能求出优化问题的解。这种方法在每一步都选择最好的选项，而不考虑可能导致最优解的所有步骤序列。在每一步都选择看起来“最好的”选项的算法称为贪心算法。一旦贪心算法求出了一个可行解，就要确定这个算法是否找到了最优解。为此，要么证明这个解是最优的，要么说明一个该算法产生非最优解的反例。为了更具体地说明这些概念，下面考虑一个用硬币找零钱的算法。

例 6 考虑用 25 美分、10 美分、5 美分和 1 美分硬币找 n 美分零钱的问题，使硬币总数尽可能少。可以设计一个找 n 美分零钱的贪心算法，在每一步都做局部最优的选择。即在每一步选择可加入已找零钱中不超过 n 美分的最大面值的硬币。例如要找 67 美分零钱，首先选择一个 25 美分硬币(剩下 42 美分)。接着选择第二个 25 美分硬币(剩下 17 美分)，随后选择一个 10 美分硬币(剩下 7 美分)，随后选择一个 5 美分硬币(剩下 2 美分)，随后选择一个 1 美分硬币(剩下 1 美分)，最后选择一个 1 美分硬币。 ■



算法 6 显示一个用一组面值的硬币找 n 美分零钱的贪心算法。

```
算法 6  找零钱贪心算法
procedure change(c_1, c_2, …, c_r: 硬币的面值, 其中 c_1>c_2>…>c_r; n: 正整数)
for i := 1 to r
  while n≥c_i
  begin
    把面值 c_i 的一个硬币加入所找零钱中
    n := n-c_i
  end
```

我们已经描述了一个用 25 美分、10 美分、5 美分和 1 美分硬币找零钱的贪心算法。证明这个算法在使用尽可能少硬币的意义下求出最优解。在开始证明之前，先说明一下存在一些硬币的组合。对于这些组合，贪心算法(算法 6)不一定得出使用尽可能少的硬币的找零方案。例如，若只有 25 美分、10 美分和 1 美分硬币(无 5 美分硬币)可用，则此贪心算法用 6 枚硬币找 30 美分零钱：1 个 25 美分和 5 个 1 美分硬币；但是本来可以用 3 个硬币，即 3 个 10 美分硬币。

引理 1 若 n 是正整数，则用 25 美分、10 美分、5 美分和 1 美分等尽可能少的硬币找出的 n 美分零钱中，至多有 2 个 10 美分、至多有 1 个 5 美分、至多有 4 个 1 美分硬币，而不能有 2 个 10 美分和 1 个 5 美分硬币。用 10 美分、5 美分和 1 美分硬币找出的零钱不能超过 24 美分。

证 用反证法。证明如果有超过规定数目的各种类型的硬币，就可以用等值的数目更少的硬币来替换。注意，如果有 3 个 10 美分硬币，就可以换成 1 个 25 美分和 1 个 5 美分硬币；如果有 2 个 5 美分硬币，就可以换成 1 个 10 美分硬币；如果有 5 个 1 美分硬币，就可以换成 1 个 5 美分硬币；如果有 2 个 10 美分和 1 个 5 美分硬币，就可以换成 1 个 25 美分硬币。由于至多可以有 2 个 10 美分、1 个 5 美分和 4 个 1 美分硬币，而不能有 2 个 10 美分和 1 个 5 美分硬币，所以当用尽可能少的硬币找 n 美分零钱时，24 美分就是用 10 美分、5 美分和 1 美分硬币能找出的最大值。 □

定理 1 贪心算法(算法 6)产生使用尽可能少的硬币找零方案。

证 将用反证法。假设存在正整数 n，使得有办法将 25 美分、10 美分、5 美分和 1 美分硬币用少于贪心算法所求出的硬币去找 n 美分零钱。首先注意，在这种找 n 美分零钱的最优方式中使用 25 美分硬币的个数 q'，一定等于贪心算法所用 25 美分硬币的个数。为说明这一点，注意贪心算法使用尽可能多的 25 美分硬币，所以 $q' \leqslant q$。但是 q' 也不能小于 q。假如 q' 小于 q，需要在这种最优方式中用 10 美分、5 美分和 1 美分硬币至少找出 25 美分零钱。而根据引理 1，这是不可能的。

由于在找零钱的这两种方式中一定有同样多的 25 美分硬币，所以在这两种方式中 10 美分、5 美分和 1 美分硬币的总值一定相等，并且这些硬币的总值不超过 24 美分。10 美分硬币的个数一定相等，因为贪心算法使用尽可能多的 10 美分硬币。而根据引理 1，当使用尽可能少的硬币找零钱时，至多使用 1 个 5 分硬币和 4 个 1 分硬币，所以在找零钱的最优方式中也使用尽可能多的 10 美分硬币。类似地，5 美分硬币的个数相等；最终，1 美分的个数相等。 □

练习

1. 列出算法 1 在表 1，8，12，9，11，2，14，5，10，4 中找最大值使用的所有步骤。
2. 判断下列过程具有和缺乏算法的哪些特征。

a)
```
procedure double (n: 正整数)
    while n>0
        n := 2n
```

b)
```
procedure divide (n: 正整数)
    while n≥0
    begin
        m := 1/n
        n := n-1
    end
```

c)
```
procedure sum(n: 正整数)
    sum := 0
    while i<10
```



$sum := sum + i$

d) **procedure** choose(a，b：整数)

x := a 或 b

3. 设计一个求表中所有整数之和的算法。
4. 描述一个算法，以 n 个整数的表作为输入，求出表中相邻整数的最大差值作为输出。
5. 描述一个算法，以 n 个按非减顺序排列的整数的表作为输入，求出所有出现两次以上的值的表。
6. 描述一个算法，以 n 个整数的表作为输入，求出表中负整数的个数。
7. 描述一个算法，以 n 个整数的表作为输入，求出表中最后一个偶数的位置，若表中没有偶数就返回 0。
8. 描述一个算法，以 n 个不同整数的表作为输入，求出表中最大偶数的位置，若表中没有偶数就返回 0。
9. 回文是从前向后读和从后向前读都一样的串。描述一个确定 n 个字符的串是否为回文的算法。
10. 设计计算 x^n 的算法，其中 x 是实数，n 是整数。（提示：首先给出一个 n 为非负整数时从 1 开始不断乘以 x 来计算 x^n 的过程。然后扩展这一过程。当 n 为负数时，利用 $x^{-n}=1/x^n$ 的事实。）
11. 描述一个交换变量 x 和 y 的值的算法，只许使用赋值。至少需要多少个赋值语句才能完成交换？
12. 描述一个只使用赋值语句用三元组(y，z，x)来代替(x，y，z)的算法。最少需要多少个赋值语句？
13. 列出在序列 1，3，4，5，6，8，9，11 中搜索 9 的所有步骤，使用的算法是：

    a)线性搜索。　　　　b)二分搜索。
14. 列出在练习 13 给出的序列中搜索 7 使用的所有步骤。
15. 给出一个算法，把整数 x 插入按增序排列的整数表 a_1，a_2，…，a_n 中合适的位置。
16. 给出一个求有限自然数序列中最小整数的算法。
17. 给出算法求有限整数表中最大元素首次出现的位置，表中的整数不一定互不相同。
18. 给出算法，求有限整数列表中最小元素最后出现的位置，列表中的整数不一定互不相同。
19. 给出一个算法计算三个整数之集合的最大值、中间值、平均值和最小值。（整数集合的中间值是把这些整数按增序排列时中间元素的值。整数集合的平均值是这些整数之和除以整数个数。）
20. 给一个求有限整数序列中最大和最小整数的算法。
21. 给出算法把任意长整数序列的头三项排成增序。
22. 给出算法求英文句子中最长的单词。（单词指字母串，句子指用空格分隔的一串单词。）
23. 给出一个算法，判断从一个有限集合到另一个有限集合的函数是否映上。
24. 给出一个算法，判断从一个有限集合到另一个有限集合的函数是否一对一。
25. 给出一个算法，逐一检查位串中每个字位是否为 1，数一数其中 1 字位的个数。
26. 改动算法 3，使二分搜索过程在每一阶段都比较 x 和 a_m，并在 $x=a_m$ 时终止。这一改动后的二分搜索算法有何优越之处？
27. 三分搜索算法为元素在增序整数表中定位，定位方法是一次次把表分成长度相等(或尽可能相等)的三个子表，每次都把搜索限制在一个合适的子表中。描述这一算法的步骤。
28. 描述一下为元素在增序整数表中定位的搜索算法，搜索方式是每次把表分成相等(或尽可能相等)的 4 个子表，并把搜索限制在一个合适的子表中。
29. 整数表的一个众数是表中出现次数不比任何别的元素出现次数少的元素。设计一个算法，求非减序整数表的一个众数。
30. 设计一个算法，求非减序整数表的所有众数(练习 29 中定义的)。
31. 设计一个算法，求整数序列中第一个与序列中排在它前面的某项相等的项。
32. 设计一个算法，找出有限整数序列中所有那些大于它前面各项之和的项。
33. 设计一个算法，求正整数序列中第一个小于其前项的项。
34. 用冒泡排序来排序 6，2，3，1，5，4，说明在每步所获得的表。
35. 用冒泡排序来排序 3，1，5，7，4，说明在每步所获得的表。
36. 用冒泡排序来排序 d，f，k，m，a，b，说明在每步所获得的表。

*37. 修改冒泡排序算法，当不再需要交换时，算法停止。用伪码表达这个更有效的算法。

38. 用插入排序来排序练习 34 中的表，说明在每步所获得的表。

39. 用插入排序来排序练习 35 中的表，说明在每步所获得的表。

40. 用插入排序来排序练习 36 中的表，说明在每步所获得的表。

选择排序首先找出表中的最小元素。把这个元素移到前面。然后找出剩余元素里的最小元素并且把它放到第二个位置。重复这个过程，直到整个表都已经排序为止。

41. 用选择排序来排序下面的表。

a)3，5，4，1，2　　b)5，4，3，2，1　　c)1，2，3，4，5

42. 用伪码写出选择排序算法。

43. 描述一个基于线性搜索的算法，确定在已经排序的表中插入一个新元素的正确位置。

44. 描述一个基于二分搜索的算法，确定在已经排序的表中插入一个新元素的正确位置。

45. 用插入排序对表 1，2，…，n 排序需要多少次比较？

46. 用插入排序对表 n，$n-1$，…，2，1 排序需要多少次比较？

二分插入排序是插入排序的一个变种，使用二分搜索技术(参见练习 44)，而不是线性搜索技术，把第 i 个元素插入到已经排序的元素中的正确位置。

47. 列出二分插入排序对表 3，2，4，5，1，6 排序使用的所有步骤。

48. 比较插入排序和二分插入排序对表 7，4，3，8，1，5，4，2 排序所用的比较次数。

*49. 用伪代码写出二分插入排序算法。

50. a)设计插入排序的一个变种，用线性搜索技术把第 j 个元素插入正确位置，即首先比较第 j 个元素与第 $j-1$ 个元素，如有必要再与第 $j-2$ 个元素比较，依次进行下去。

b)用上述算法排序 3，2，4，5，1，6。

c)用这个算法解练习 45。

d)用这个算法解练习 46。

51. 当一个元素表接近于正确顺序时，用插入排序或者用练习 50 描述的变种，哪一种更好？

52. 采用贪心算法，用 25 美分、10 美分、5 美分和 1 美分硬币找出下列零钱：

a)87 美分　　b)49 美分　　c)99 美分　　d)33 美分

53. 采用贪心算法，用 25 美分、10 美分、5 美分和 1 美分硬币找出下列零钱：

a)51 美分　　b)69 美分　　c)76 美分　　d)60 美分

54. 采用贪心算法，用 25 美分、10 美分和 1 美分(但是无 5 美分)硬币找出练习 52 中的各种零钱。对于哪些零钱数，贪心算法使用尽可能少的这些面值的硬币？

55. 采用贪心算法，用 25 美分、10 美分和 1 美分(但是无 5 美分)硬币找出练习 53 中的各种零钱。对于哪些零钱数，贪心算法使用尽可能少的这些面值的硬币？

56. 证明若有面值 12 美分的硬币，则用 25 美分、12 美分、10 美分、5 美分和 1 美分硬币的贪心算法，不一定总是用尽可能少的硬币找零钱。

57. a)描述一个贪心算法，解决在报告厅安排一组候选演讲中的部分演讲的问题，在每一步都选择与已经安排的演讲不冲突的结束时间最早的候选演讲。(在 3.3 节将证明这个贪心算法总是得出演讲数尽可能多的安排。)

b)用(a)部分的算法从一组候选演讲中选择，在报告厅安排尽可能多的演讲。假设这些演讲的开始和结束时间(均为上午)是：9:00 和 9:45；9:30 和 10:00；9:50 和 10:15；10:00 和 10:30；10:10 和 10:25；10:30 和 10:55；10:15 和 10:45；10:30 和 11:00；10:45 和 11:30；10:55 和 11:25；11:00 和 11:15。

2.2 函数的增长

2.2.1 引言

2.1 节讨论了算法的概念及解决各式各样问题的算法，包括在表中搜索元素和对表进行排序。在 2.3 节将研究这些算法使用的操作步数。特别是，将估计线性和二分搜索算法在 n 个元素的序列中搜索元素所用的比较次数。还将估计冒泡排序和插入排序对 n 个元素的表进行排序所用的比较次数。解决一个问题所需的时间不仅取决于所用的操作步数，还取决于运行实现算法的程序所用的硬件和软件。但是，当改变实现算法所用的硬件和软件时，给从前所需的时间乘以一个常数，就可以精确地逼近解决规模为 n 的问题所需的时间。例如，在一台超级计算机上解规模为 n 的问题可能比在一台个人电脑上快 1 百万倍。而这个 1 百万的因子并不取决于 n（也许会有一点点的依赖关系）。使用本节介绍的大 O 记号的一个好处，就是可以估计函数增长而不用担心常数因子或低阶项。这意味着使用大 O 记号，不用担心实现算法所用的硬件和软件。另外，使用大 O 记号，可以假设算法使用的不同运算都花费相等的时间，这大大简化了分析。

大 O 记号广泛用于估计当输入增长时一个算法所用的运算数。有了这个记号，可以确定当输入规模增大时，用一个特定算法来解题是否实际可行。另外，使用大 O 记号，可以比较两个算法，以判断当输入规模增大时哪个算法更有效。例如，如果有两个解题的算法，一个使用 $100n^2+17n+4$ 步运算，另一个使用 n^3 步运算，大 O 记号可以帮助我们看出，当 n 很大时，第一个算法使用的运算少得多，即使对于小的 n 值，比如 $n=10$，第一个算法使用的运算较多。

本节介绍大 O 记号和相关的大 Ω 和大 Θ 记号，将解释如何构造大 O、大 Ω 和大 Θ 估计，并证明对算法分析所用到的一些重要函数的估计。

2.2.2 大 O 记号

我们常用一个专门的符号描述函数的增长，下面的定义 1 描述这一符号。

定义 1 令 f 和 g 为从整数集合或实数集合到实数集合的函数，如果有常数 C 和 k，使得只要 $x>k$，就有

$$|f(x)|\leqslant C|g(x)|$$

我们说 $f(x)$ 是 $O(g(x))$（$O(g(x))$ 读作"$f(x)$ 是大 O $g(x)$"）。

大 O 记号定义中的常数 C 和 k 称为"$f(x)$ 是 $O(g(x))$"这一关系的见证。为了证明 $f(x)$ 是 $O(g(x))$，只需要这一关系的一对见证。换句话说，要证明 $f(x)$ 是 $O(g(x))$，只需要找出一对常数 C 和 k（即见证），使得只要 $x>k$ 就有 $|f(x)|\leqslant C|g(x)|$。

注意，当有一对"$f(x)$ 是 $O(g(x))$"这一关系的见证时，就有*无穷多对的见证*。要看明白这一点，注意若 C 和 k 是一对见证，那么任意一对 C' 和 k'（其中 $C<C'$ 和 $k<k'$）也是一对见证，因为只要 $x>k'>k$，就有 $|f(x)|\leqslant C|g(x)|\leqslant C'|g(x)|$。

求一对见证的一种有用方法是：首先选择 k 的值，要求当 $x>k$ 时容易估计 $|f(x)|$ 的大小，并看看能否用这个估计求 C 的值，要求对于 $x>k$ 有 $|f(x)|<C|g(x)|$，如例 1。

例 1 证明 $f(x)=x^2+2x+1$ 是 $O(x^2)$。

解 注意当 $x>1$ 时容易估计 $f(x)$ 的大小，因为当 $x>1$ 时 $x<x^2$ 且 $1<x^2$，所以只要 $x>1$ 就有

$$0 \leqslant x^2+2x+1 \leqslant x^2+2x^2+x^2=4x^2$$

如图 2-2 所示。因此可以取 $C=4$ 和 $k=1$ 做见证证明 $f(x)$是 $O(x^2)$。即只要 $x>1$ 就有 $f(x)=x^2+2x+1<4x^2$。（注意这里不必用绝对值符号，因为当 x 为正数时，等式中所有函数都是正的。）

另一办法是，当 $x>2$ 时估计 $f(x)$的大小。当 $x>2$ 时 $2x \leqslant x^2$ 且 $1 \leqslant x^2$。于是如果 $x>2$，就有

$$0 \leqslant x^2+2x+1 \leqslant x^2+x^2+x^2=3x^2$$

所以 $C=3$ 和 $k=2$ 也是“$f(x)$是 $O(x^2)$”这一关系的见证。

注意在 $f(x)$是 $O(x^2)$这一关系中，x^2 可以被函数值大于 x^2 的任何函数取代。例如，$f(x)$是 $O(x^3)$，$f(x)$是 $O(x^2+2x+7)$等。

x^2 是 $O(x^2+2x+1)$也成立，因为只要 $x>1$ 就有 $x^2<x^2+2x+1$。这意味着 $C=1$ 和 $k=1$ 是 $f(x)$是 $O(x^2+2x+1)$这一关系的见证。■

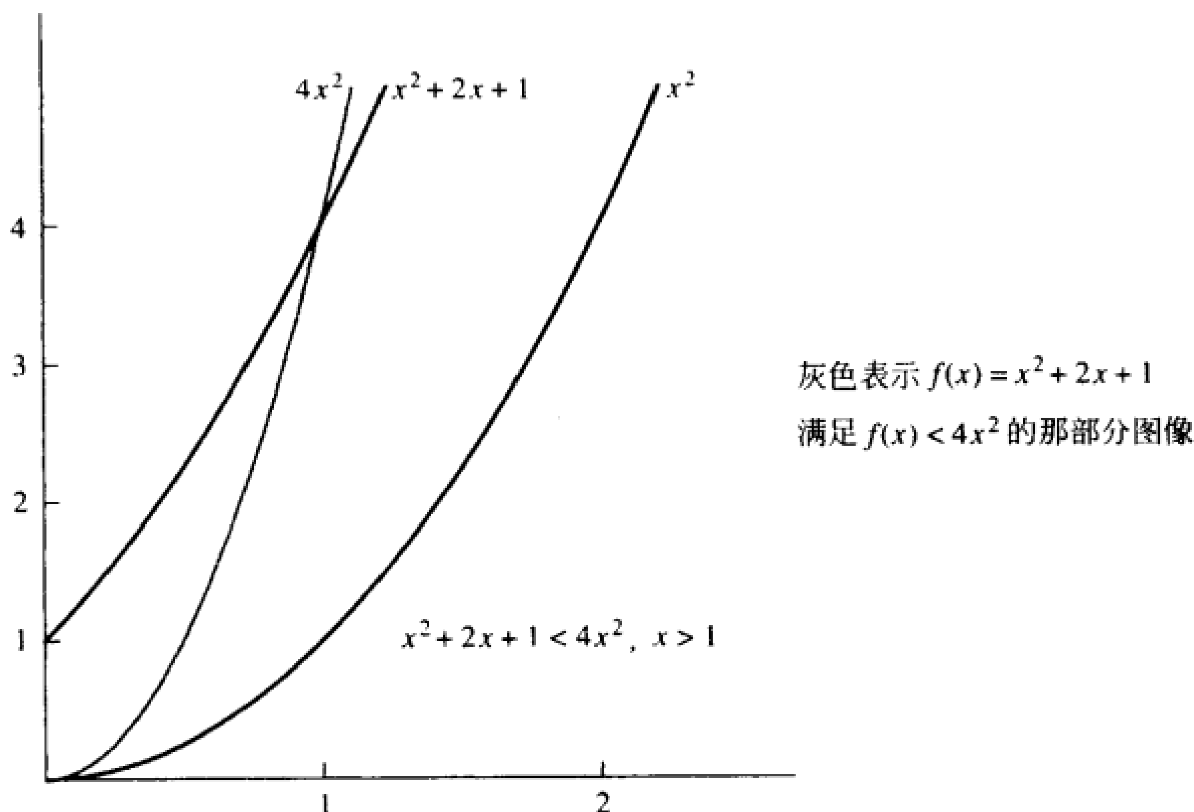


图 2-2 函数 x^2+2x+1 是 $O(x^2)$

注意，在例 1 中有两个函数 $f(x)=x^2+2x+1$ 和 $g(x)=x^2$，使得 $f(x)$是 $O(g(x))$而且$g(x)$是 $O(f(x))$。后一事实可以从不等式 $x^2 \leqslant x^2+2x+1$ 得到，这个不等式对所有非负实数 x 都成立。我们把满足上述两个大 O 关系的函数 $f(x)$和 $g(x)$称为同阶的。本节后面还要讨论这个概念。

注意 $f(x)$是 $O(g(x))$的事实有时写作 $f(x)=O(g(x))$。不过这一写法中的等号并不代表真正的相等。其实对于这些函数的定义域中足够大的数而言，函数 f 和 g 的值之间有个不等式成立。

大 O 记号已经在数学中使用了差不多一个世纪。在 2.3 节会看到，这一记号在计算机科学

中也广泛应用于算法分析。德国数学家 Paul Bachmann㊀1892 年在他的一本重要数论著作中首先引入了大 O 记号。大 O 记号有时又以德国数学家朗道(Edmund Landau)㊁的名字命名为*朗道符号*。朗道在他的著作中始终使用这一符号。大 O 记号在计算机科学中的普遍使用归功于高德纳(Donald Knuth)㊂，他还引入了将在本节随后介绍的大 Ω 和大 Θ 记号。

当 $f(x)$ 是 $O(g(x))$，而对足够大的 x，$h(x)$ 是一个绝对值大于 $g(x)$ 的函数时，有 $f(x)$ 是 $O(h(x))$。换言之，在 $f(x)$ 是 $O(g(x))$ 这一关系中，可以用具有更大绝对值的函数替换函数 $g(x)$。要看清这一点，注意如果

$$|f(x)| \leqslant C|g(x)| \quad 若\ x>k$$

并且如果对所有 $x>k$，有 $|h(x)|>|g(x)|$，那么

$$|f(x)| \leqslant C|h(x)| \quad 若\ x>k$$

于是 $f(x)$ 是 $O(h(x))$。

㊀ **巴赫曼(Paul Gustav Heinrich Bachmann，1837—1920)**是路德教牧师的儿子，继承了他父亲虔诚的生活方式和对音乐的热爱。尽管巴赫曼早期的数学学习并不顺利，他的一位老师还是发现了他的数学才能。在瑞士从肺结核的病痛中康复以后，巴赫曼学习了数学。首先在柏林大学，随后在哥廷根大学。在哥大他听了著名数论家迪里赫里(Dirichlet)的课。1862 年在德国数论家库默尔(Kummer)指导下获博士学位；他的论文内容是群论。巴赫曼曾先后担任布来斯劳(Breslau)大学和明斯特(Münster)大学教授。从教授位置上退休后，他继续数学写作、弹钢琴并且在报纸上作音乐评论。巴赫曼的数学论著包括 5 卷本对数论结果与方法的评述，2 卷本的初等数论，一本关于无理数的书和一本关于以费尔马最后定理闻名的数论猜想的书。在他 1892 年的书《*Analytische Zahlentheorie*》(解析数论)中引入了大 O 记号。

㊁ **朗道(Edmund Landau，1877—1938)**是一位柏林妇科医生的儿子，在柏林完成高中和大学教育。1899 年在佛罗本尼乌斯(Frobenius)的指导下获得博士学位。朗道首先在柏林大学任教，后到哥廷根大学，在那里任教授直到纳粹迫使他离开。朗道对数学的贡献主要在解析数论领域，特别是就素数分布得到几个重要的结论。他完成了三卷本的数论评注及关于数论和数学分析的其他书籍。

㊂ **高德纳(Donald E. Knuth，1938 年生)**在密尔沃基长大。他父亲在那里的路德高中教授簿记，并拥有一家小型的印刷厂。高德纳是个优秀的学生，多次获得学业成就奖。他以反传统的方式运用才能，在八年级时赢了一次找字比赛，成绩是用“Ziegler's Giant Bar”中的字母组成了 4500 个单词。这次胜利为他的母校赢得一台电视，为班上的同学每人赢了一根棒糖。

高德纳在 Case Institute of Technology (开思理工学院)上学时就放弃音乐而主修物理，做了艰难的决策。然后他又从物理转为数学，1960 年获学士学位。由于教师们认为他的工作杰出，同时以特别奖的形式授予他硕士学位。在开思，他管理篮球队，并用他的才能发明了一个估价每位球员价值的公式。这一新奇的方法被《*Newsweek*》和 CBS 电视网的 Walter Cronkite 报道。1960 年高德纳开始在加州理工学院做研究生，并于 1963 年获博士学位。在这期间他还担任顾问，为不同的计算机写编译程序。

高德纳 1963 年加入了加州理工学院的教师队伍，一直到 1968 年担任斯坦福大学教授。1992 年为集中精力写作，他作为荣誉退休教授。他特别感兴趣的是为他的《计算机程序设计艺术》丛书完成新卷写作并更新旧卷。这套丛书是 1962 年他还是研究生时以编译程序为中心开始写作的，至今已对计算机科学的发展产生了意义深远的影响。在常用的行话中，“高德纳”就意味着《计算机程序设计艺术》，也就意味着数据结构和算法这一类问题的答案。

高德纳是现代计算复杂度研究的奠基人。他对编译程序做出了基础性的贡献。对数学印刷的不满激发他发明了现在广泛使用的 TeX 和 Metafont 系统。TeX 已经成为计算机排印的一个标准语言。他获得的众多奖项包括 1974 年的图灵奖和卡特总统授给他的 1979 年国家技术奖。

高德纳为计算机科学和数学方面内容广泛的专业期刊写了许多文章。不过他的头一篇作品是 1957 年还是一年级新生时写的，是对称为“The Potrzebie Systems of Weights and Measures”(重量和长度的 Potrzebie 系统)的计量系统的模仿小品。该文首先登在 MAD 杂志，已经多次重印。他与父亲一样是一位教堂风琴手，还为风琴作曲。高德纳相信，可以像写诗或作曲一样编写计算机程序。

对他书中的每个错误，高德纳付 2.56 美元给第一个发现的人，对每个有意义的建议，他付 0.32 美元。如果你寄给他一封信指出一个错误(你只能寄普通信件，因为他已放弃阅读电子邮件)，他最终会通知你，你是否是第一个告诉他这一错误的人。要耐心等待，因为他收到的邮件太多。(作者寄给高德纳一封报告错误的信，几年以后才收到回信，他告诉我，我的报告比首先报告这一错误的信晚到了好几个月。)

在使用大 O 记号时，在 $f(x)$是$O(g(x))$这一关系中函数 g 总是选得尽可能小。(有时选自某个参考函数集合，例如形为 x^n 的函数，其中 n 为正整数。)

在随后的讨论中，差不多总是处理只有正值的函数。对这样的函数来说，在用大 O 做估计时可以不必涉及绝对值。图 2-3 说明 $f(x)$是$O(g(x))$的关系。

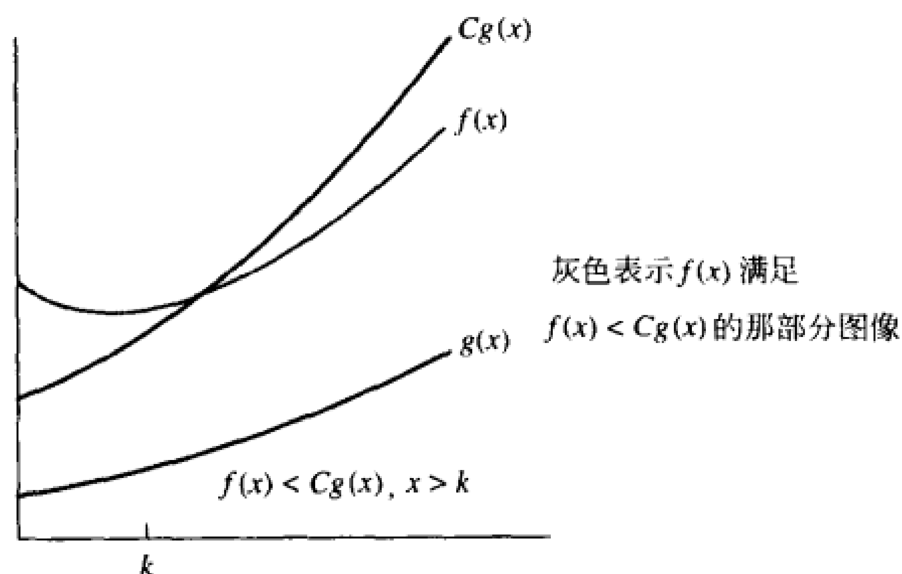


图 2-3　函数 $f(x)$是$O(g(x))$

例 2 说明如何用大 O 记号来估计函数的增长。

例 2　证明 $7x^2$ 是 $O(x^3)$。

解　注意当 $x>7$ 时，就有 $7x^2<x^3$。(在 $x>7$ 两边乘以 x^2 得出这个不等式。)因此，取$C=1$ 和 $k=7$ 做见证，证明 $7x^2$ 是 $O(x^3)$这一关系。另一种方法是，当 $x>1$ 时，有 $7x^2<7x^3$，于是 $C=7$ 和 $k=1$ 也是 $7x^2$ 是 $O(x^3)$这一关系的见证。 ■

例 3　例 2 证明了 $7x^2$ 是 $O(x^3)$。x^3 也是 $O(7x^2)$吗?

解　要判定 x^3 是否 $O(7x^2)$，必须判断是否存在常数 C 和 k，使得只要 $x>k$，就有 $x^3\leqslant C(7x^2)$。这一不等式等价于不等式 $x\leqslant 7C$，这是在上一不等式两边除以 x^2 得到的。注意不可能存在这样的 C，无论 k 是什么，对所有 $x>k$ 有 $x\leqslant 7C$，因为 x 可以任意大。于是 x^3 不是 $O(7x^2)$。 ■

2.2.3　一些重要的大 O 结果

多项式常用于估计函数的增长，与其每当多项式出现时都要分析多项式增长，不如找一个总是可以估计多项式增长的结论。下面的定理给出这种结论。这个定理证明了多项式的首项支配其增长，断定 n 次及低次多项式是 $O(x^n)$。

定理 1　令 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$，其中 a_0，a_1，…，a_{n-1}，a_n 为实数，那么 $f(x)$是 $O(x^n)$。

证　用三角不等式(参见 1.5 节练习 35)，如果 $x>1$，就有

$$\begin{aligned}|f(x)|&=|a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0|\\&\leqslant|a_n|x^n+|a_{n-1}|x^{n-1}+\cdots+|a_1|x+|a_0|\\&=x^n(|a_n|+|a_{n-1}|/x+\cdots+|a_1|/x^{n-1}+|a_0|/x^n)\\&\leqslant x^n(|a_n|+|a_{n-1}|+\cdots+|a_1|+|a_0|)\end{aligned}$$

这说明只要 $x>1$，就有

$$|f(x)| \leqslant Cx^n$$

其中 $C=|a_n|+|a_{n-1}|+\cdots+|a_1|+|a_0|$。于是 $f(x)$ 是 $O(x^n)$。 □

现在举几个与定义域为正整数集合的函数有关的例子。

例 4 怎样用大 O 记号估计前 n 个正整数的和？

解 由于前 n 个正整数都不超过 n，所以

$$1+2+\cdots+n \leqslant n+n+\cdots+n = n^2$$

由此不等式知 $1+2+\cdots+n$ 是 $O(n^2)$，取 $C=1$ 和 $k=1$ 作为见证。（本例大 O 关系中的函数定义域为正整数集合。） ■

例 5 中用大 O 估计阶乘函数及其对数函数。这些估计对分析排序过程中使用的步数有重要作用。

例 5 给出阶乘函数和阶乘函数之对数的大 O 估计，其中阶乘函数 $f(n)=n!$ 的定义为：只要 n 是正整数，

$$n! = 1 \cdot 2 \cdot \cdots \cdot n$$

而 $0!=1$。例如

$$1! = 1, \quad 2! = 1 \cdot 2 = 2, \quad 3! = 1 \cdot 2 \cdot 3 = 6, \quad 4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24$$

注意函数 $n!$ 增长迅速。例如

$$20! = 2\,432\,902\,008\,176\,640\,000$$

解 注意到乘积中的每一项都不超过 n，就能得到 $n!$ 的大 O 估计。于是

$$\begin{aligned} n! &= 1 \cdot 2 \cdot 3 \cdot \cdots \cdot n \\ &\leqslant n \cdot n \cdot n \cdot \cdots \cdot n \\ &= n^n \end{aligned}$$

这一不等式说明 $n!$ 是 $O(n^n)$。对用于估计 $n!$ 的不等式两边同时取对数，得

$$\log n! \leqslant \log n^n = n\log n$$

这表明 $\log n!$ 是 $O(n\log n)$，取 $C=1$ 和 $k=1$ 作为见证。 ■

例 6 在 3.3 节将证明对正整数 n，$n<2^n$，用这一不等式可以得知 n 是 $O(2^n)$。（取 $k=C=1$ 做见证。）由于对数函数是增函数，只要在这一不等式两边取对数（以 2 为底），得

$$\log n < n$$

于是 $\log n$ 是 $O(n)$。（仍取 $k=C=1$ 做见证。）

如果以 b 为底取对数，其中 b 不等于 2，就有 $\log_b n$ 是 $O(n)$，因为

$$\log_b n = \frac{\log n}{\log b} < \frac{n}{\log b}$$

其中 n 是正整数。取 $C=1/\log b$ 和 $k=1$ 做见证。（用了附录 A 定理 3 的公式 $\log_b n=\log n/\log b$。） ■

前面提过，大 O 符号常用于估计一个特定的计算机过程或算法解题时需要的操作数。做这些估计时常用的函数包括：

$$1, \log n, n, n\log n, n^2, 2^n, n!$$

用微积分可以证明，上述每个函数都小于表中列在随后的函数，准确地说，每个函数与随后的函数的比在 n 无限增长时趋向于 0。图 2-4 给出了这些函数的图像，图中函数值的每个刻度都是前面的刻度的两倍。

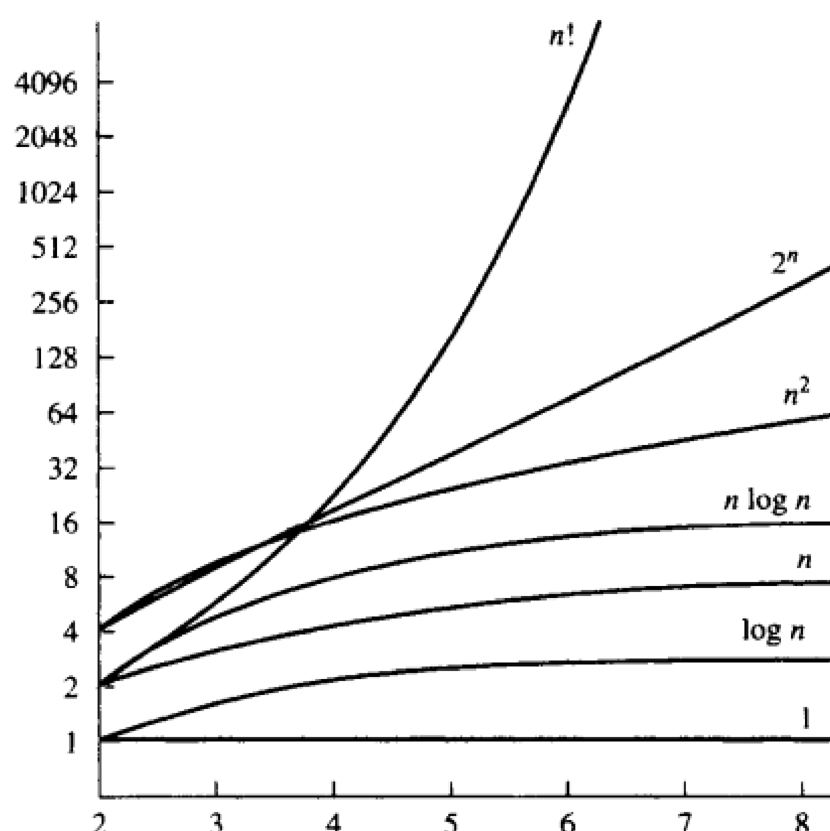


图 2-4 大 O 估计常用函数增长图

2.2.4 函数组合的增长

许多算法都由两个或更多独立的子过程组成，使用这样的算法，用计算机求解一定大小的问题所需要的步数是这些过程使用的步数的和。用大 O 估计需要的步数，必须找出对每个子过程所用步数的估计，再把这些估计组合在一起。

只要能把不同的大 O 估计组合在一起，就能提供对函数组合的大 O 估计。特别是往往要估计两个函数之和与之积的增长。如果已知两个函数各自的大 O 估计，能得到什么结论呢？要弄清两个函数之和与之积有什么样的估计，假定 $f_1(x)$ 是 $O(g_1(x))$，$f_2(x)$ 是 $O(g_2(x))$。

从大 O 记号的定义，有常数 C_1、C_2、k_1 和 k_2，使得：当 $x>k_1$ 时

$$|f_1(x)|\leqslant C_1|g_1(x)|$$

当 $x>k_2$ 时，

$$|f_2(x)|\leqslant C_2|g_2(x)|$$

要估计 $f_1(x)$ 与 $f_2(x)$ 之和，注意

$$\begin{aligned}|(f_1+f_2)(x)|&=|f_1(x)+f_2(x)|\\&\leqslant|f_1(x)|+|f_2(x)| \qquad (\text{用三角不等式 } |a+b|\leqslant|a|+|b|)\end{aligned}$$

当 x 大于 k_1 和 k_2 时，从 $|f_1(x)|$ 和 $|f_2(x)|$ 的不等式得：

$$\begin{aligned}|f_1(x)|+|f_2(x)|&\leqslant C_1|g_1(x)|+C_2|g_2(x)|\\&\leqslant C_1|g(x)|+C_2|g(x)|\\&=(C_1+C_2)|g(x)|\\&=C|g(x)|\end{aligned}$$

其中 $C=C_1+C_2$，而 $g(x)=\max(|g_1(x)|, |g_2(x)|)$。（这里 $\max(a, b)$ 表示 a 和 b 的最大值，亦即 a 和 b 较大的一个。）

这一不等式说明 $|(f_1+f_2)(x)\leqslant C|g(x)|$ 在 $x>k$ 时成立，其中 $k=\max(k_1, k_2)$。把这一有用的结果作为定理 2。

定理 2 假定 $f_1(x)$是$O(g_1(x))$，$f_2(x)$是$O(g_2(x))$，那么$(f_1+f_2)(x)$是$O(\max(|g_1(x)|, |g_2(x)|))$。

对 f_1 和 f_2 的大 O 估计往往会使用同样的函数 g。在此情况下，由于 $\max(g_1(x), g_2(x))=g(x)$，定理 2 说明$(f_1+f_2)(x)$也是$O(g(x))$，下面的推论 1 就是这一结果。

推论 1 假定 $f_1(x)$和 $f_2(x)$都是$O(g(x))$，那么$(f_1+f_2)(x)$也是$O(g(x))$。

用类似的方法可以推导出 f_1 和 f_2 的乘积的大 O 估计。当 x 大于 $\max(k_1, k_2)$时，有

$$\begin{aligned}|(f_1f_2)(x)| &= |f_1(x)||f_2(x)| \\ &\leqslant C_1|g_1(x)|C_2|g_2(x)| \\ &\leqslant C_1C_2|(g_1g_2)(x)| \\ &\leqslant C|(g_1g_2)(x)|\end{aligned}$$

其中 $C=C_1C_2$。从这一不等式知 $f_1(x)f_2(x)$是$O(g_1g_2)$，因为存在常数 C 和 k，即 $C=C_1C_2$ 和 $k=\max(k_1, k_2)$，使得只要 $x>k$ 就有 $|(f_1f_2)(x)|\leqslant C|g_1(x)g_2(x)|$。定理 3 即是这一结果。

定理 3 假定 $f_1(x)$是$O(g_1(x))$，$f_2(x)$是$O(g_2(x))$，那么$(f_1f_2)(x)$是$O(g_1(x)g_2(x))$。

用大 O 符号估计函数之目的，是选一个相对增长较慢的函数 $g(x)$，使得 $f(x)$是$O(g(x))$。下面的例子说明怎样利用定理 2 和定理 3 来实现这一目标。下面的几个例子中给出的这一类分析常用于分析用计算机程序解题时所用的时间。

例 7 给出 $f(n)=3n\log(n!)+(n^2+3)\log n$ 的大 O 估计，其中 n 是一个正整数。

解 首先估计乘积 $3n\log(n!)$，从例 5 知道，$\log(n!)$是$O(n\log n)$。由这一估计及 $3n$ 是$O(n)$的事实，定理 3 给出的估计为 $3n\log(n!)$是 $O(n^2\log n)$。

下一步估计乘积$(n^2+3)\log n$。由于$(n^2+3)<2n^2$ 在 $n>2$ 时成立，则 n^2+3 是 $O(n^2)$。因此，由定理 3 知$(n^2+3)\log n$ 是 $O(n^2\log n)$。用定理 2 把两个乘积的估计组合在一起，得 $f(n)=3n\log(n!)+(n^2+3)\log n$ 是 $O(n^2\log n)$。■

例 8 给出 $f(x)=(x+1)\log(x^2+1)+3x^2$ 的大 O 估计。

解 首先找$(x+1)\log(x^2+1)$的大 O 估计。注意$(x+1)$是 $O(x)$。另外当 $x>1$ 时 $x^2+1\leqslant 2x^2$。于是当 $x>2$ 时，

$$\log(x^2+1)\leqslant\log(2x^2)=\log 2+\log x^2=\log 2+2\log x\leqslant 3\log x$$

这说明 $\log(x^2+1)$是 $O(\log x)$。

从定理 3 知$(x+1)\log(x^2+1)$是 $O(x\log x)$，由于 $3x^2$ 是 $O(x^2)$，定理 2 说明 $f(x)$是 $O(\max(x\log x, x^2))$。当 $x>1$ 时，由于 $x\log x\leqslant x^2$，于是 $f(x)$是 $O(x^2)$。■

2.2.5 大 Ω 与大 Θ 记号

大 O 记号广泛用于描述函数的增长，但有局限性。特别是当 $f(x)$是 $O(g(x))$时，我们只有用 $g(x)$估计 $f(x)$的大小对大 x 值的一个上限。大 O 记号不能提供对大 x 值的 $f(x)$之大小的下限。为此使用大 Ω 记号。当希望给出函数 $f(x)$的规模相对于参照函数 $g(x)$的上限和下限时，使用大 Θ 符号。大 Ω 和大 Θ 符号都是高德纳在 1970 年引入的，他引入这两个符号的动机是，纠正人们需要函数的上限和下限时对大 O 符号的误用。

现在定义大 Ω 符号并解释其用法。然后再定义大 Θ 并解释其用法。

在大 O 和大 Ω 记号之间有很强的联系，特别是 $f(x)$是 $\Omega(g(x))$当且仅当 $g(x)$是 $O(f(x))$。这一事实的证明作为练习留给读者。

定义 2 令 f 和 g 为从整数集合或实数集合到实数集合的函数，如果存在正常数 C 和 k，使

得在 $x>k$ 时

$$|f(x)|\geqslant C|g(x)|$$

我们说 $f(x)$ 是 $\Omega(g(x))$（读作 $f(x)$ 是大 $\Omega g(x)$）。

例9 函数 $f(x)=8x^3+5x^2+7$ 是 $\Omega(g(x))$，其中 $g(x)=x^3$。由于 $f(x)=8x^3+5x^2+7\geqslant 8x^3$ 对所有正实数都成立，所以上述说法成立。这等价于 $g(x)=x^3$ 是 $O(8x^3+5x^2+7)$，而只须把不等式颠倒过来写直接就得到这一结论。■

通常，重要的是知道相对一个简单参照函数某函数增长的阶。所谓简单函数，指的是如 x^n（其中 n 是正整数）或 c^x（其中 $c>1$）这样的函数。要知道函数增长的阶，就需要了解函数大小的上界和下界。也就是说，给定一个函数 $f(x)$，我们想要一个参照函数 $g(x)$，使得 $f(x)$ 是 $O(g(x))$ 和 $f(x)$ 是 $\Omega(g(x))$。下面定义的大 Θ 记号表达这两个关系，既提供函数大小的上界，又提供函数大小的下界。

定义3 令 f 和 g 为从整数集合或实数集合到实数集合的函数，如果 $f(x)$ 是 $O(g(x))$ 及 $f(x)$ 是 $\Omega(g(x))$，就说 $f(x)$ 是 $\Theta(g(x))$。若 $f(x)$ 是 $\Theta(g(x))$，就说"$f(x)$ 是大 $\Theta g(x)$"，也说 $f(x)$ 是 $g(x)$ 阶的。

若 $f(x)$ 是 $\Theta(g(x))$，$g(x)$ 也是 $\Theta(f(x))$，在使用大 Θ 符号时，$\Theta(g(x))$ 中的 $g(x)$ 通常是一个相对简单的参照函数，如 x^n，c^x，$\log x$ 等，而 $f(x)$ 可能会相对复杂些。

例10 已证明（例4）前几个正整数的和是 $O(n^2)$，这个和是 n^2 阶的吗？

Extra Examples

解 令 $f(n)=1+2+3+\cdots+n$，由于已经是 $f(n)=O(n^2)$，为证明 $f(n)$ 是 n^2 阶的，只需找到正整数 C 使得对足够大的 n，有 $f(n)>Cn^2$。为获得这一和的下界，可以丢掉这些项中的前一半，只把大于 $\lceil n/2\rceil$ 的项加起来，得

$$\begin{aligned}1+2+\cdots+n &\geqslant \lceil n/2\rceil+(\lceil n/2\rceil+1)+\cdots+n\\ &\geqslant \lceil n/2\rceil+\lceil n/2\rceil+\cdots+\lceil n/2\rceil\\ &=(n-\lceil n/2\rceil+1)\lceil n/2\rceil\\ &\geqslant (n/2)(n/2)\\ &=n^2/4\end{aligned}$$

这说明 $f(n)$ 是 $\Omega(n^2)$，结论是：$f(n)$ 是 n^2 阶的。用符号来写，$f(n)$ 是 $\Theta(n^2)$。■

可以证明，如果能找到正实数 C_1、C_2 和正实数 k，使得对 $x>k$，

$$C_1|g(x)|\leqslant|f(x)|\leqslant C_2|g(x)|$$

则 $f(x)$ 是 $\Theta(g(x))$，因为上式说明 $f(x)$ 是 $O(g(x))$ 及 $f(x)$ 是 $\Omega(g(x))$。

例11 证明 $3x^2+8x\log x$ 是 $\Theta(x^2)$。

解 因为 $0\leqslant 8x\log x\leqslant 8x^2$，所以对 $x\geqslant 1$，$3x^2+8x\log x\leqslant 11x^2$。从而 $3x^2+8x\log x$ 是 $O(x^2)$，显然 x^2 是 $O(3x^2+8x\log x)$，也就是 $3x^2+8x\log x$ 是 $\Theta(x^2)$。■

一个有用的事实是，多项式的首项决定多项式的阶。例如，若 $f(x)=3x^5+x^4+17x^3+2$，那么 $f(x)$ 是 x^5 阶的。下面的定理说的就是这一事实，其证明留在本节末的练习中由读者来完成。

定理4 令 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$，其中 a_0，a_1，$\cdots$，a_n 为实数，且 $a_n\neq 0$。则 $f(x)$ 是 x^n 阶的。

例12 多项式 $3x^8+10x^7+221x^2+1444$，$x^{19}-18x^4-10\,112$ 和 $-x^{99}+40001x^{98}+100\,003x$ 分别是 x^8、x^{19} 和 x^{99} 阶的。■

不幸的是，正如高德纳已观察到的，大 O 记号常被粗心的作者和演讲者误用，以为其含义与大 Θ 相同。当你见到有人使用大 O 记号时，别忘了它常被误用。近来的倾向是，如果需要函

数大小的上界和下界，就使用大 Θ。

练习

1. 判断下列各函数是否为 $O(x)$。
   a) $f(x)=10$　b) $f(x)=3x+7$　c) $f(x)=x^2+x+1$
   d) $f(x)=5\log x$　e) $f(x)=\lfloor x\rfloor$　f) $f(x)=\lceil x/2\rceil$
2. 判断下列各函数是否为 $O(x^2)$。
   a) $f(x)=17x+11$　b) $f(x)=x^2+1000$　c) $f(x)=x\log x$
   d) $f(x)=x^4/2$　e) $f(x)=2^x$　f) $f(x)=\lfloor x\rfloor\cdot\lceil x\rceil$
3. 用 $f(x)$是 $O(g(x))$这一事实的定义证明 x^4+9x^3+4x+7 是 $O(x^4)$。
4. 用 $f(x)$是 $O(g(x))$是一事实的定义证明 2^x+17 是 $O(3^x)$。
5. 证明$(x^2+1)/(x+1)$是 $O(x)$。
6. 证明$(x^3+2x)/(2x+1)$是 $O(x^2)$。
7. 对下列每个函数求最小的整数 n，使 $f(x)$是 $O(x^n)$。
   a) $f(x)=2x^3+x^2\log x$　b) $f(x)=3x^3+(\log x)^4$
   c) $f(x)=(x^4+x^2+1)/(x^3+1)$　d) $f(x)=(x^4+5\log x)/(x^4+1)$
8. 对下列每个函数求最小的整数 n，使 $f(x)$是 $O(x^n)$。
   a) $f(x)=2x^2+x^3\log x$　b) $f(x)=3x^5+(\log x)^4$
   c) $f(x)=(x^4+x^2+1)/(x^4+1)$　d) $f(x)=(x^3+5\log x)/(x^4+1)$
9. 证明 $x^2+4x+17$ 是 $O(x^3)$，但 x^3 不是 $O(x^2+4x+17)$。
10. 证明 x^3 是 $O(x^4)$，但 x^4 不是 $O(x^3)$。
11. 证明 $3x^4+1$ 是 $O(x^4/2)$，而且 $x^4/2$ 也是 $O(3x^4+1)$。
12. 证明 $x\log x$ 是 $O(x^2)$但 x^2 不是 $O(x\log x)$。
13. 证明 2^n 是 $O(3^n)$但 3^n 不是 $O(2^n)$。
14. 若 $g(x)$是下列函数，x^3 是 $O(g(x))$吗？(例如，如果 $g(x)=x+1$，这一问题问的是 x^3 是 $O(x+1)$吗？)
    a) $g(x)=x^2$　b) $g(x)=x^3$　c) $g(x)=x^2+x^3$
    d) $g(x)=x^2+x^4$　e) $g(x)=3^x$　f) $g(x)=x^3/2$
15. 解释一个函数是 $O(1)$的含义。
16. 证明若 $f(x)$是 $O(x)$，那么 $f(x)$是 $O(x^2)$。
17. 假定 $f(x)$，$g(x)$和 $h(x)$为函数，使 $f(x)$是 $O(g(x))$，$g(x)$是 $O(h(x))$。证明 $f(x)$是 $O(h(x))$。
18. 令 k 为正整数，证明 $1^k+2^k+\cdots+n^k$ 是 $O(n^{k+1})$。
19. 对下列各函数给出一个尽可能好的大 O 估计。
    a) $(n^2+8)(n+1)$　b) $(n\log n+n^2)(n^3+2)$　c) $(n!+2^n)(n^3+\log(n^2+1))$
20. 给出下列各函数的大 O 估计。在你的估计 $f(x)$是 $O(g(x))$中，要使用阶最小的一个简单函数。
    a) $(n^3+n^2\log n)(\log n+1)+(17\log n+19)(n^3+2)$
    b) $(2^n+n^2)(n^3+3^n)$　c) $(n^n+n2^n+5^n)(n!+5^n)$
21. 给下列各函数一个大 O 估计，在你的估计 $f(x)$是 $O(g(x))$中，要使用阶最小的一个简单函数。
    a) $n\log(n^2+1)+n^2\log n$　b) $(n\log n+1)^2+(\log n+1)(n^2+1)$
    c) $n^{2^n}+n^{n^2}$
22. 对练习 1 中的各函数，判断它是否为 $\Omega(x)$和 $\Theta(x)$。
23. 对练习 2 中的各函数，判断它是否为 $\Omega(x^2)$和 $\Theta(x^2)$。
24. a) 证明 $3x+7$ 是 $\Theta(x)$。

b)证明 $2x^2+x-7$ 是 $\Theta(x^2)$。 c)证明 $\lfloor x+1/2 \rfloor$ 是 $\Theta(x)$。

d)证明 $\log(x^2+1)$ 是 $\Theta(\log_2 x)$。 e)证明 $\log_{10} x$ 是 $\Theta(\log_2 x)$。

25. 证明 $f(x)$ 是 $\Theta(g(x))$，当且仅当 $f(x)$ 是 $O(g(x))$ 且 $g(x)$ 是 $O(f(x))$。
26. 若 $f(x)$ 和 $g(x)$ 是从实数集到实数集的函数，求证 $f(x)$ 是 $O(g(x))$，当且仅当 $g(x)$ 是 $\Omega(f(x))$。
27. 若 $f(x)$ 和 $g(x)$ 是从实数集到实数集的函数，求证 $f(x)$ 是 $\Theta(g(x))$，当且仅当存在正常数 k，C_1 和 C_2 使得只要 $x>k$，就有 $C_1|g(x)| \leqslant |f(x)| \leqslant C_2|g(x)|$。
28. a)找出练习 27 中要求的 k，C_1 和 C_2，直接证明 $3x^2+x+1$ 是 $\Theta(3x^2)$。

    b)用图像表示(a)中的关系，即给出函数 $3x^2+x+1$、$C_1 \cdot 3x^2$ 和 $C_2 \cdot 3x^2$ 的图像，在 x 轴上标出 k，其中 k，C_1，C_2 是(a)中你找到的证明 $3x^2+x+1$ 是 $\Theta(3x^2)$ 的常数。
29. 用图像表示 $f(x)$ 是 $\Theta(g(x))$ 这一关系。画出 $f(x)$、$C_1|g(x)|$、$C_2|g(x)|$ 的图像，并在 x 轴上标出常数 k。
30. 解释函数为 $\Omega(1)$ 的含义。
31. 解释函数为 $\Theta(1)$ 的含义。
32. 给出前 n 个奇正整数之乘积的一个大 O 估计。
33. 若 f 和 g 为实数值函数，使 $f(x)$ 是 $O(g(x))$，证明 $f^k(x)$ 是 $O(g^k(x))$。(注意 $f^k(x)=f(x)^k$。)
34. 若 $f(x)$ 是 $O(\log_b x)$，其中 $b>1$，证明 $f(x)$ 是 $O(\log_a x)$，其中 $a>1$。
35. 若 $f(x)$ 是 $O(g(x))$，其中 f 和 g 是无限增长的函数，证明 $\log|f(x)|$ 是 $O(\log|g(x)|)$。
36. 假定 $f(x)$ 是 $O(g(x))$，能否推断 $2^{f(x)}$ 是 $O(2^{g(x)})$？
37. 令 $f_1(x)$ 和 $f_2(x)$ 为从实数集合到正实数集合的函数，证明：若 $f_1(x)$ 和 $f_2(x)$ 均为 $\Theta(g(x))$，其中 $g(x)$ 是从实数集到正实数集的一个函数，则 $f_1(x)+f_2(x)$ 是 $\Theta(g(x))$。如果 $f_1(x)$ 能取负值。这一结论还成立吗？
38. 假定 $f(x)$，$g(x)$ 和 $h(x)$ 是函数，使 $f(x)$ 是 $\Theta(g(x))$，$g(x)$ 是 $\Theta(h(x))$。证明 $f(x)$ 是 $\Theta(h(x))$。
39. 若 $f_1(x)$，$f_2(x)$ 为从正整数集合到正实数集合的函数，且 $f_1(x)$ 和 $f_2(x)$ 都是 $\Theta(g(x))$。是否 $(f_1-f_2)(x)$ 也是 $\Theta(g(x))$？证明它成立或给出一个反例。
40. 若 $f_1(x)$ 和 $f_2(x)$ 为从正整数集合到实数集合的函数，且 $f_1(x)$ 是 $\Theta(g_1(x))$，$f_2(x)$ 是 $\Theta(g_2(x))$。证明 $(f_1 f_2)(x)$ 是 $\Theta(g_1 g_2(x))$。
41. 找从正整数集合到实数集合的函数 f 和 g，使 $f(n)$ 不是 $O(g(n))$，$g(n)$ 也不是 $O(f(n))$。
42. 用图像表示，$f(x)$ 是 $\Omega(g(x))$。画出函数 $f(x)$ 和 $C g(x)$，同时在 x 轴上标出常数 k。
43. 若 $f_1(x)$ 是 $\Theta(g_1(x))$，$f_2(x)$ 是 $\Theta(g_2(x))$，且对所有实数 $x>0$，$f_2(x)(0$，$g_2(x)\neq 0$，证明 $(f_1/f_2)(x)$ 是 $\Theta((g_1/g_2)(x))$。
44. 证明：若 $f(x)=a_n x^n+a_{n-1}x^{n-1}+\cdots+a_1 x+a_0$，其中 a_0，a_1，…，a_n 为实数，且 $a_n\neq 0$，则 $f(x)$ 是 $\Theta(x^n)$。

大 O、大 Θ 和大 Ω 等符号可以推广到多元函数，例如，语句 $f(x, y)$ 是 $O(g(x, y))$ 的含义是：存在常数 C、k_1 和 k_2，使得对 $x>k_1$ 和 $y>k_2$，$|f(x, y)| \leqslant C|g(x, y)|$。

45. 定义语句 $f(x, y)$ 是 $\Theta(g(x, y))$。
46. 定义语句 $f(x, y)$ 是 $\Omega(g(x, y))$。
47. 证明 $(x^2+xy+x\log y)^3$ 是 $O(x^6 y^3)$。
48. 证明 $x^5 y^3+x^4 y^4+x^3 y^5$ 是 $\Omega(x^3 y^3)$。
49. 证明 $\lfloor xy \rfloor$ 是 $O(xy)$。
50. 证明 $\lceil xy \rceil \Omega(xy)$。

以下的问题处理另一类渐近记号，称为小 o 记号，由于小 o 记号以极限概念为基础，故微积分知识是必要的。如果

$$\lim_{x\to\infty}\frac{f(x)}{g(x)}=0$$

就说 $f(x)$是 $o(g(x))$（读作 $f(x)$是小 o $g(x)$）。

51. （需要微积分）求证

a) x^2 是 $o(x^3)$　b) $x\log x$ 是 $o(x^2)$　c) x^2 是 $o(2^x)$　d) x^2+x+1 不是 $o(x^2)$

52. （需要微积分）

a)求证：若函数 $f(x)$和 $g(x)$使得 $f(x)$是 $o(g(x))$，c 为常数，则 $cf(x)$是 $o(g(x))$，其中$(cf)(x)=cf(x)$。

b)求证：若 $f_1(x)$，$f_2(x)$和 $g(x)$为函数，使得 $f_1(x)$是 $o(g(x))$，$f_2(x)$是 $o(g(x))$，则$(f_1+f_2)(x)$是 $o(g(x))$，其中$(f_1+f_2)(x)=f_1(x)+f_2(x)$。

53. （需要微积分）画出 $x\log x$，x^2 及 $x\log x/x^2$ 的图像以表示 $x\log x$ 是 $o(x^2)$。解释这一图像怎样说明 $x\log x$ 是 $o(x^2)$。

54. （需要微积分）用图像表示 $f(x)$是 $O(g(x))$。画出 $f(x)$，$g(x)$和 $f(x)/g(x)$的图像。

*55. （需要微积分）假定 $f(x)$是 $o(g(x))$。能否由此推出 $2^{f(x)}$ 是 $o(2^{g(x)})$？

*56. （需要微积分）假定 $f(x)$是 $o(g(x))$。能否由此推出 $\log|f(x)|$ 是 $o(\log|g(x)|)$？

57. （需要微积分）本题中的两部分描述了小 o 和大 O 符号之间的关系。

a)求证：若函数 $f(x)$和 $g(x)$使 $f(x)$是 $o(g(x))$，则 $f(x)$是 $O(g(x))$。

b)求证：若函数 $f(x)$和 $g(x)$使 $f(x)$是 $O(g(x))$，那么不一定能推出 $f(x)$是 $o(g(x))$。

58. （需要微积分）求证：若 $f(x)$是 n 阶多项式，$g(x)$是 m 阶多项式，且 $m>n$，则 $f(x)$是 $o(g(x))$。

59. （需要微积分）求证：若 $f_1(x)$是 $O(g(x))$，$f_2(x)$是 $o(g(x))$，那么 $f_1(x)+f_2(x)$是 $O(g(x))$

60. （需要微积分）令 H_n 为第 n 调和数

$$H_n=1+\frac{1}{2}+\frac{1}{3}+\cdots+\frac{1}{n}$$

求证 H_n 是 $O(\log n)$。（提示：建立不等式

$$\sum_{j=2}^{n}\frac{1}{j}<\int_1^n\frac{1}{x}\mathrm{d}x$$

办法是证明对 $j=2,3,\cdots,n$，以 $j-1$ 到 j 为底、$1/j$ 为高的所有这些长方形的面积之和小于曲线 $y=1/x$ 下面从 2 到 n 的这一面积。）

*61. 求证 $n\log n$ 是 $O(\log n!)$。

☞62. 判断 $\log(n!)$是 $\Theta(n\log n)$是否成立，给出理由。

令 $f(x)$和 $g(x)$为从实数集合到实数集合的函数。如果 $\lim_{x\to\infty}(f(x)/g(x))=1$，就说 $f(x)$和 $g(x)$是渐近的，写作 $f(x)\sim g(x)$。

63. （需要微积分）对下列每对函数，判断 f 和 g 是否渐近。

a) $f(x)=x^2+3x+7$，$g(x)=x^2+10$

b) $f(x)=x^2\log x$，$g(x)=x^3$

c) $f(x)=x^4+\log(3x^8+7)$，$g(x)=(x^2+17x+3)^2$

d) $f(x)=(x^3+x^2+x+1)^4$，$g(x)=(x^4+x^3+x^2+x+1)^3$

e) $f(x)=\log(x^2+1)$，$g(x)=\log x$

f) $f(x)=2^{x+3}$，$g(x)=2^{x+7}$

g) $f(x)=2^{2^x}$，$g(x)=2^{x^2}$

2.3 算法的复杂度

2.3.1 引言

什么条件下算法能提供令人满意的解？首先，算法必须给出正确的答案。第 3 章将讨论如何说明算法的正确性。其次，算法必须有效率，本节讨论算法的效率。

怎样分析算法的效率呢？一种度量是当输入值具有一定规模时，计算机按此算法解题所花的时间。第二种度量是输入值具有一定规模时，实现这一算法计算机需要多大内存。

这些问题都涉及算法的计算复杂度。分析解决特定规模的问题需要的时间是算法的时间复杂度。分析需要的计算机内存是算法的空间复杂度。在实现一个算法时，时间和空间复杂度的考虑都是实质性的。显然，了解算法是一微秒、一分钟还是几亿年才能给出答案是很重要的。类似地，必须能提供所需的内存才能解题，所以必须考虑空间复杂度。

空间复杂度的考虑与实现算法时使用的特定数据结构紧密相关。由于本书对数据结构不做详细讨论，所以不考虑空间复杂度，而将注意力集中在时间复杂度上。

2.3.2 时间复杂度

在输入规模一定的情况下，算法的时间复杂度可用算法使用的运算次数表示。度量时间复杂度使用的运算包括：整数比较，整数加法，整数乘法，整数除法或任何其他基本运算。

时间复杂度是用运算次数而不是计算机实际使用的时间来表示，因为在执行基本运算时不同的计算机需要的时间不同。把所有运算分解成计算机使用的基本位运算是相当复杂的。而且，现存最快的计算机执行位运算(例如两个位的加、乘、比较或交换)的时间是 10^{-9} 秒(1 纳秒)，但个人电脑可能需要 10^{-6} 秒(1 微秒)，做同样的运算时间相差 1000 倍。

我们用 2.1 节求有限整数集合最大值的算法 1 来说明怎样分析算法的时间复杂度。

例 1 描述 2.1 节求集合最大元素的算法 1 的时间复杂度。

解 由于比较是该算法使用的基本运算，就以比较的次数作为其时间复杂度的度量。

要求以任意顺序列出的 n 个元素之集合的最大元素，首先让临时最大值等于表中的初始项。然后在比较后判断是否未到表的终点，将临时最大值与第二项比较，如果第二项大，就用第二项的值更新临时最大值。这一过程继续下去，对表中的每一项都进行两次比较：一次判断是否未到表终点，另一次判断是否需要更新临时最大值。由于对从第二项到第 n 项的每一项用两次比较，再加上一次在 $i=n+1$ 时退出循环的比较，所以应用这一算法时使用的比较次数恰为 $2(n-1)+1=2n-1$。因此，求 n 元集合最大值的算法的时间复杂度是 $\Theta(n)$，这是用算法使用的比较次数来度量的。 ■

下面将分析搜索算法的时间复杂度。

例 2 描述线性搜索的时间复杂度。

解 该算法使用的比较次数将用来度量时间复杂度。算法中每次循环都做两次比较。一次判断是否已到表的终点，一次比较 x 和表的一个项。最后还要在循环以外再做一次比较。于是当 $x=a_i$ 时，最多需要做 $2i+1$ 次比较。当 x 不在表中时，最多需要 $2n+2$ 次比较。在这种情况下，$2n$ 次比较用来确定 x 不是 a_i，$i=1, 2, \cdots, n$，再加上一次比较用于退出循环，和一次循环外的比较。所以当 x 不在表中时，共用 $2n+2$ 次比较。从而线性搜索最多需要 $O(n)$次比较。 ■

最坏情形复杂度 例 2 中做的这类复杂度分析是最坏情形分析。算法的最坏情形指的是把算法应用于一定规模的问题最多需要多少次运算。最坏情形分析说明算法需要多少次运算就保证给出问题的解答。

例 3 描述二分搜索算法的时间复杂度。

解 为简化描述，假定表 $a_1, a_2, \cdots, a_n$ 中有 $n=2^k$ 个元素，其中 k 是非负整数。注意 $k=\log n$。(如果表中元素个数 n 不是 2 的幂，那么这个表可以看作一个有 2^{k+1} 个元素的大表的一部分，其中 $2^k<n<2^{k+1}$。因此 2^{k+1} 是大于 n 的 2 的最小幂。)

在算法的每一阶段，都要比较 i 和 j 来判断待搜索的表是否包含 1 个以上的元素，其中 i 和 j 分别是当前待搜索表的第一项和最后项的位置。如果 $i<j$，就要做一次比较，判断 x 是否大于待搜索表的中间元素。

在第一阶段将搜索限于含 2^{k-1} 个元素的表。至此已使用了两次比较，这一过程继续下去，每一阶段都用两次比较把搜索限制在长度减半的表中，换言之，第一阶段在表中含 2^k 个元素时，用两次比较把表减半，在表中含 2^{k-1} 个元素时又用两次比较，在表长减为 2^{k-2} 时又用两次比较，如此等等，直到两次比较后，表长为 $2^1=2$。最后当表中只剩一个元素时，再用一次比较确定没有其他元素留在表中，还要一次比较判断这一项是否为 x。

因此，在表中有 2^k 个元素时，二分搜索需要 $2k+2=2\log n+2$ 次比较。(如果 n 不是 2 的幂，最初的表扩展为含 2^{k+1} 个项的表，其中 $k=\lfloor \log n\rfloor$，搜索需要最多 $2\lceil \log n\rceil+2$ 次比较。)于是二分搜索算法需要最多 $\Theta(\log n)$ 次比较。由此分析知，在最坏的情况下，二分搜索算法比线性搜索效率高。 ■

平均情形复杂度 除最坏情形分析以外，还有称为平均情形分析的另一类重要的复杂度分析。这类分析是就一定规模的问题，对所有输入求解使用的平均运算次数。平均情形时间复杂度分析一般比最坏情形分析复杂得多。不过，线性搜索算法的平均情形分析不难完成，例 4 就是个说明。

例 4 描述线性搜索算法的平均情形性能，假定元素 x 在表中。

解 已知 x 在表中时，有 n 类可能的输入。如果 x 是表的第一项，需要比较 3 次：一次判断是否已到表终点，一次比较 x 和第一项，再在循环外比较一次。如果 x 是表的第二项，还要增加 2 次比较，所以总共要比较 5 次。一般来说，若 x 是表的第 i 项，i 次循环中的每一次都要做 2 次比较，外加循环外一次，所以共需要 $2i+1$ 次比较。于是平均比较次数是：

$$\frac{3+5+7+\cdots+(2n+1)}{n}=\frac{2(1+2+3+\cdots+n)+n}{n}$$

在 3.3 节将证明

$$1+2+3+\cdots+n=\frac{n(n+1)}{2}$$

所以，线性搜索使用的平均比较次数(已知 x 在表中)是

$$\frac{2(n(n+1)/2)+n}{n}=n+2$$

这是 $\Theta(n)$。 ■

注意 第 4 章的分析中假定了 x 在被搜索的表中，而且 x 在一切位置的机会都相等。当 x 可能不在表中时，也可以对这一算法做平均情形分析(参看本节末的练习 13)。

注意 尽管把判断是否到达循环结尾所需的比较也算进来，但通常这些比较是不算进来的。从现在起忽略这些比较。

两个排序算法的最坏情形复杂度 例 5 和例 6 分析冒泡排序和插入排序的最坏情形复杂度。

例 5 用比较次数来衡量，冒泡排序的最坏情形复杂度是什么?

解 冒泡排序(其描述在 2.1 节例 4)通过一遍遍处理一个表来对这个表排序。在每一遍冒泡排序都一次次比较相邻元素，必要时交换相邻元素。在第 i 遍开始时，$i-1$ 个最大元素保证是在正确位置上。在这一遍，使用 $n-i$ 次比较，所以排序 n 个元素的表时，冒泡排序所使用的总比较次数是

$$(n-1)+(n-2)+\cdots+2+1=\frac{(n-1)n}{2}$$

以上利用了将在 3.3 节证明的求和公式。注意冒泡排序总是使用这么多次的比较，因为即使在某个中间步骤表已经完全排好了，冒泡排序仍继续进行。于是冒泡排序使用 $n(n-1)/2$ 次比较，所以用比较次数来衡量，冒泡排序的最坏情形复杂度是 $\Theta(n^2)$。 ■

例 6 用比较次数来衡量，插入排序的最坏情形复杂度是什么？

解 插入排序(其描述在 2.1 节例 5)把第 j 个元素插入到前 $j-1$ 个已排好顺序的元素中的正确位置上。插入排序用线性搜索技术来做到这一点，依次比较第 j 个元素与后续各项，直到找到大于或等于这个元素的一项或者比较 a_j 与它自身为止，因为 a_j 不小于它自身。于是在最坏情形下，把第 j 个元素插入正确位置需要 j 次比较。所以排序 n 个元素的表时，插入排序所使用的总比较次数是

$$2+3+\cdots+n=\frac{n(n+1)}{2}-1$$

以上利用了将在 3.3 节证明的求连续整数之和的公式，并且注意和中缺少第一项 1。注意，如果较小的元素起初是在表的尾部，则插入排序可能使用少得多的比较。所以插入排序的最坏情形复杂度是 $\Theta(n^2)$。 ■

2.3.3 理解算法的复杂度

表 2-1 中给出了描述算法时间复杂度的几个常用术语。例如，一个求 n 个元素的表前 100 项中最大项的算法对前 100 项的序列应用算法 1，其中 n 是满足 $n\geqslant 100$ 的整数。这个算法具有*常数复杂度*，因为无论 n 是什么值，这个算法都使用 99 次比较(读者可以验证)。线性搜索算法具有(最坏情形或平均情形)*线性复杂度*，而二分搜索算法具有(最坏情形)*对数复杂度*。许多重要的算法都具有 $n\log n$ 复杂度，例如将在第 3 章学习的归并排序。

如果一个算法有复杂度 $O(n^b)$，其中 b 是满足 $b\geqslant 1$ 的整数，那么这个算法有*多项式复杂度*。例如，冒泡排序算法是多项式时间算法，因为这个算法在最坏情形下使用 $O(n^2)$ 次比较。如果一个算法有时间复杂度 $O(b^n)$，其中 $b>1$，就说这个算法有*指数复杂度*。为确定 n 个变量的复合命题是否可满足而检查变量的所有可能赋值的算法是指数复杂度算法，因为这个算法用 $O(2^n)$ 次运算。最后，如果一个算法有时间复杂度 $O(n!)$，那么这个算法具有*阶乘复杂度*。求推销员可以用来访问 n 个城市的所有顺序的算法具有阶乘复杂度，第 3 章将讨论这个算法。

表 2-1 算法复杂度常用术语

复杂度	术 语	复杂度	术 语
$O(1)$	常数复杂度	$O(n^b)$	多项式复杂度
$O(\log n)$	对数复杂度	$O(b^n)$，$b>1$	指数复杂度
$O(n)$	线性复杂度	$O(n!)$	阶乘复杂度
$O(n\log n)$	$n\log n$ 复杂度		

能用具有多项式最坏情形复杂度的算法解决的问题称为*易解的*，因为只要问题的规模合理，就可期望算法在相对短的时间内给出解答。不过，如果在大 O 估计中的多项式次数高(如 100 次)，或如果多项式的系数特别大，算法都可能会花特别长的时间来解题。所以，能用具有多项式最坏情形复杂度的算法来解决的问题，即使对于相对较小的输入值，也不能保证能在合理时间内得到解答。幸运的是，实践中这种估计用到的多项式之次数和系数都不大。

对不能用最坏情形多项式时间复杂度的算法解决的问题，情况要糟得多。这种问题称为难

解的。一般来说，需要特别大量的时间来解决最坏情形的问题，既使输入值很小时也是这样。当然也许会有例外。不过以最坏情形分析其复杂度的算法在解决实际问题时，大多数情况都比最坏情形快得多。如果允许少量情况下问题不能在合理的时间内得到解答，那么平均情形时间复杂度是对算法解题时间的更好度量。工业上许多重要的问题都认为是难解的，但实践中，对日常生活中出现的所有输入基本上都能得到问题的解。另一种处理实践中出现的不易处理的问题的方法是，不求精确解，而以近似解代替。也许存在求近似解的快速算法，甚至还能保证这些近似解和精确解相差不太大。

甚至存在一些可以证明没有解题算法存在的问题，这种问题称为*不可解的*(相对于可以用某个算法求解的可解问题而言)。第一个证明存在不可解问题的是伟大的英国数学家和计算机科学家 Alan Turing(图灵)。他证明不可解的是*停机问题*。这个问题的输入是一个程序，加上对这一程序的输入，问题要问的是当这个程序以给定的输入执行时会不会停止。在 3.1 节将学习停机问题。(第 11 章有图灵小传及对他某些其他工作的介绍。)

算法复杂度的研究远超出在这里所能介绍的。要注意，人们相信许多可解的问题都没有多项式最坏情形时间复杂度的算法能解答，但是一旦有了一个解答，却可以用多项式时间来验证。能以多项式时间验证解的问题属于 NP 类㊀(易解的问题属于 P 类)。还有称为 *NP 完全问题*的重要的一类问题，这类问题的性质是，只要其中任何一个问题能用一个多项式时间最坏情形算法来解答，那么所有这些问题都能用多项式时间最坏情形算法解答。

可满足性问题是 NP 完全问题的一个例子——可以快速地验证复合命题的一组变量赋值让这个命题成真，但是至今没有找到求这种成真赋值的多项式时间算法。(例如，穷举搜索所有可能的赋值需要 $\Theta(2^n)$ 次位运算，其中 n 是复合命题的变量数。)另外，如果知道了一个解答可满足性问题的多项式时间算法，那么所有已知在这个问题类中的问题就都有多项式时间算法(而在这个类中有许多重要问题)。

尽管人们做了广泛的研究，并没有找到多项式时间最坏情形的算法能解决这类问题中的任何一个。虽然没有证明，但人们普遍接受的说法是，没有 NP 完全问题能用多项式时间解决。要更多了解算法复杂度的信息，参阅本书最后为本节列出的文献，包括[CoLeRi 90]。

注意，算法时间复杂度的大 O 估计表达的是解题需要的时间如何随输入数据规模的增长而改变。实践中使用的是能得到证明的最好估计(即参照函数最小的估计)。不过时间复杂度的大 O 估计不能直接翻译成计算机使用的实际时间量。一个原因是，大 O 估计 $f(n)$ 是 $O(g(n))$ 表示的是：有常数 C 和 k，使得在 $n>k$ 时 $f(n)\leqslant Cg(n)$，其中 $f(n)$ 是算法的时间复杂度，$g(n)$ 是参照函数。所以不知道不等式中的常数 C 和 k，就不能用这一估计来判定算法使用的运算次数的上界。此外前文已经说过，每次运算所需要的时间还因运算类型不同和使用的计算机不同而有差异。(还要注意，算法时间复杂度的大 O 估计只能对算法需要的最坏情形时间以输入值规模的函数的形式提供上界，而不提供下界。要提供下界，就应使用大 Θ 函数估计。不过为简单起见，在讨论算法时间复杂度时，将用大 O 估计，同时要懂得大 Θ 估计能提供更多的信息。)

不过，如果算法使用的所有运算都能归结为计算机使用的位运算，那么算法解答一定规模的问题所需的时间是能够确定的。表 2-2 给出的是算法解答各种规模的问题需要的时间，算法使用的位运算次数如表中所示。需要的时间超过 10^{100} 年时，表中用星号表示。(在 2.5 节将讨论整数加法和乘法使用的位运算次数。)在构造这一表格时，假定了每次位运算需要的时间是 10^{-9} 秒，这是今天最快的计算机需要的时间，将来有了更快的计算机，这些时间将会减少。

㊀ NP 表示非确定型多项式(nondeterministic polynomial)时间。

表 2-2 算法使用的计算机时间

问题规模	使用的位运算					
n	$\log n$	n	$n\log n$	n^2	2^n	$n!$
10	3×10^{-9}s	10^{-8}s	3×10^{-8}s	10^{-7}s	10^{-6}s	3×10^{-3}s
10^2	7×10^{-9}s	10^{-7}s	7×10^{-7}s	10^{-5}s	4×10^{13}年	*
10^3	1.0×10^{-8}s	10^{-6}s	1×10^{-5}s	10^{-3}s	*	*
10^4	1.3×10^{-8}s	10^{-5}s	1×10^{-4}s	10^{-1}s	*	*
10^5	1.7×10^{-8}s	10^{-4}s	2×10^{-3}s	10s	*	*
10^6	2×10^{-8}s	10^{-3}s	2×10^{-2}s	17min	*	*

了解计算机需要多长时间来解题很重要。例如，如果算法需要 10 小时，也许值得花费这些机时(和金钱)来解决问题。但如果算法需要百亿年来解题，就没有理由消耗资源来实现这一算法。现代技术最有趣的现象之一是计算机速度和内存空间的大量增长。减少计算机解题时间的另一重要方式是并行处理，这是同时执行几个运算序列的技术。

有效算法，包括大多数多项式时间复杂度的算法，从重大技术进步中得到的好处最大。但这些技术进步对于克服指数或阶乘时间复杂度的算法的复杂度没有什么帮助。由于计算速度的增加，计算机内存的增加，再加上使用能做并行处理的算法，5 年前认为无法解的问题现在可以当做日常事务处理了，而且可以肯定这句话在 5 年以后仍然成立。

练习

1. 用 2.1 节练习 16 给出的算法求含 n 个自然数的序列中的最小自然数，要使用多少次比较？
2. 写一个算法，把任意长的表中前 4 项排成增序，并证明以算法使用的比较次数度量其时间复杂度是$O(1)$。
3. 假定已知一个元素是含 32 个元素的一个表的前 4 个元素之一。线性搜索或二分搜索会不会较快地为此元素定位？
4. 给出计算 x^{2^k} 使用的乘法次数，假定从 x 开始逐次做平方(求 x^2，x^4 等)。这样计算 x^{2^k} 是否比每次都直接用 x 乘，直到合适的次数效率高？
5. 给出下述算法所使用的比较次数的大 O 估计，这一算法检查位串的每一位是否为 1 以计算位串中 1 的个数(参看 2.1 节练习 25)。

*6. a)证明下面的算法给出的是位串 S 中 1 的个数。

```
procedure bit count (S: 位串)
count := 0
while S≠0
begin
   count := count+1
   S := S∧(S−1)
end{count 的值是 S 中 1 的个数}
```

其中 $S-1$ 是把 S 中最右边的 1 改为 0，同时把这一位右边的所有 0 均改为 1 得到的位串。(回忆一下，$S\wedge(S-1)$是 S 和 $S-1$ 的按位合取运算。)

b)上述算法要做多少次按位合取运算才能求出位串 S 中 1 的个数？

7. 计算多项式 $a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 在 $x=c$ 的值的传统算法可以用伪码表示为：

```
procedure polynomial(c, a0, a1, …, an: 实数)
power := 1
```

```
y := a_0
for i=1 to n
begin
    power := power * c
    y := y + a_i * power
end {y=a_n c^n + a_{n-1} c^{n-1} + ⋯ + a_1 c + a_0}
```

其中 y 的最终值即是该多项式在 $x=c$ 的值。

a)按上述算法一步步计算 $3x^2+x+1$ 在 $x=2$ 的值。

b)准确给出计算 n 阶多项式在 $x=c$ 的值使用的乘法和加法的次数。(不要计算增加循环变量的值所做的加法。)

8. 有一个算法计算多项式的值比上道练习中给出的传统算法效率高(以使用的乘法和加法次数来度量)，这个算法称为*霍纳法*。下面的伪码说明怎样用这一方法计算 $a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 在 $x=c$ 的值。

```
procedure Horner (c, a_0 a_1, ⋯, a_n: 实数)
y := a_n
for i=1 to n
y := y * x + a_{n-i}
{y=a_n c^n + a_{n-1} c^{n-1} + ⋯ + a_1 c + a_0}
```

a)按上述算法一步步计算 $3x^2+x+1$ 在 $x=2$ 的值。

b)准确给出此算法计算 n 阶多项式在 $x=c$ 的值使用的乘法和加法的次数。(不要计算增加循环变量的值所做的加法。)

9. 需使用 $f(n)$ 次位运算的算法在 1 秒钟内能解决多大的问题，假定每次位运算均能在 10^{-9} 秒完成？设 $f(n)$ 的值如下：

a) $\log n$　　b) n　　c) $n\log n$　　d) n^2　　e) 2^n　　f) $n!$

10. 假定算法解答规模为 n 的问题需 $2n^2+2^n$ 次位运算，每次位运算需 10^{-9} 秒，对于下面给出的 n 值，计算算法解题花的时间。

a)10　　b)20　　c)50　　d)100

11. 假定每次位运算需要的时间如下，使用 2^{50} 次位运算的算法需要多少时间？

a) 10^{-6} s　　b) 10^{-9} s　　c) 10^{-12} s

12. 判定下述情况需要的最少比较次数，也就是最好情况性能。

a)用 2.1 节的算法 1 求 n 个整数的序列的最大值。

b)用线性搜索在 n 个元素的表中为一个元素定位。

c)用二分搜索在 n 个元素的表中为一个元素定位。

13. 分析线性搜索的平均情形性能，假定恰有一半的情况 x 不在表中；而且当 x 在表中时，它出现在表的任何位置的可能性都一样。

14. 对一个算法和它涉及的某种运算来说，如果不存在能用更少的此种运算解同一问题的其他算法，就说该算法对此种运算是*最优的*。

a)证明 2.1 节算法 1 对于整数比较的次数而言是最优的。(*注意*：不考虑用于循环管理的比较。)

b)对于整数比较次数而言，线性搜索是最优的吗？(不计循环管理用法的比较。)

15. 描述 2.1 节练习 27 给出的三分搜索算法用比较次数度量的最坏情形时间复杂度。

16. 描述 2.1 节练习 28 中给出的搜索算法用比较次数度量的最坏情形时间复杂度。

17. 分析你为 2.1 节练习 29 设计的算法的最坏情形时间复杂度，该算法为非减序整数表的一个众数定位。

18. 分析你为 2.1 节练习 30 设计的算法的最坏情形时间复杂度，该算法为非减序整数表所有众数定位。

19. 分析你为2.1节练习31设计的算法的最坏情形时间复杂度，该算法求整数序列中第一个与它前面某项相等的项。
20. 分析你为2.1节练习32设计的算法的最坏情形时间复杂度，该算法求序列中所有那些大于其前面各项之和的项。
21. 分析你为2.1节练习33设计的算法的最坏情形时间复杂度，该算法求序列中第一个小于前一项的项。
22. 用比较次数来度量，确定2.1节练习5的算法的最坏情形时间复杂度，该算法求整数排序表中所有多次出现的值。
23. 用比较次数来度量，确定2.1节练习9的算法的最坏情形时间复杂度，该算法判断一个串是否回文。
24. 选择排序(参见2.1节练习41前的说明)给 n 个项排序要用多少次比较？用选择排序的比较次数来度量，给出选择排序复杂度的大 O 估计。
25. 对于在2.1节练习47前的短文中描述的二分插入排序，用比较次数和交换的项数来度量，找出最坏情形复杂度的大 O 估计。
26. 证明用比较次数来度量，用25美分、10美分、5美分和1美分硬币找 n 美分零钱的贪心算法有 $O(n)$ 复杂度。
27. 描述一下用下列算法搜索一个表中的元素时，当表的规模从 n 加倍到 $2n$ 时，其中 n 是正整数，这些算法在最坏情形下所用的比较次数如何变化。

    a)线性搜索　　b)二分搜索
28. 描述一下使用下列排序算法时，当要排序的表的规模从 n 加倍到 $2n$ 时，其中 n 是正整数，这些算法在最坏情形下所用的比较次数如何变化。

    a)冒泡排序　　b)插入排序

    c)选择排序(见2.1节练习41前的说明)　　d)二分插入排序(见2.1节练习47前的说明)

2.4 整数和除法

2.4.1 引言

离散数学中涉及整数及其性质的这一部分属于数学中称为*数论*的分支。本节是介绍数论的三节中的第一节。本节先要回顾一下数论中的某些基本概念，包括整除性、最大公约数和同余算术。在2.5节将介绍数论中的几个重要的算法，并把2.1节和2.3节中关于算法及其复杂度的内容与本节引入的概念结合在一起。例如，将介绍求两个正整数的最大公约数的算法及用二进制表示做计算机算术的算法。最后在2.6节继续学习数论，介绍一些重要的结果以及这些结果在计算机算术和密码学中的应用(密码学研究信息保密)。

本节将要展开的内容以整除性为基础。从整除性建立的一个重要概念是*素数*。素数是大于1且只能被1和它自己整除的整数。判断一个整数是否为素数对密码学应用来说很重要。数论中的一个重要的定理，即*算术基本定理*的结论是每个正整数都能唯一地分解为素数的乘积。把整数分解为素数因子在密码学中是很重要的。用一个正整数除一个整数得到商和余数。贯穿计算机科学的同余算术处理的就是这些余数。本节将讨论同余算术的三个应用：生成伪随机数、为文件分配内存地址和为信息加密和解密。

2.4.2 除法

当一个整数被第二个整数除的时候，商可能是也可能不是整数。例如，$12/3=4$ 是整数，而 $11/4=2.75$ 不是。这样就有了下面的定义。

定义1 如果 a 和 b 是整数，$a\neq0$，若有整数 c 使 $b=ac$，就说 a 整除 b。在 a 整除 b 时，a

是 b 的一个因子，b 是 a 的倍数。符号 $a\mid b$ 表示 a 整除 b。当 a 不能整除 b 时写成 $a\nmid b$。

注意 可以用量词把 $a\mid b$ 表示成 $\exists c(ac=b)$，其中论域是整数集合。

图 2-5 中的数轴显示的是哪些整数能被正整数 d 整除。

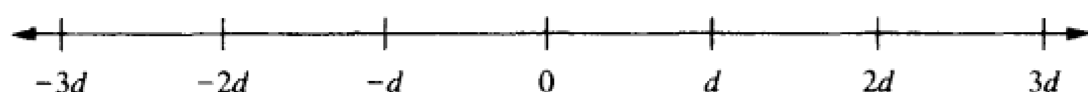


图 2-5 被正整数 d 整除的整数

例 1 判断是否 $3\mid 7$ 和 $3\mid 12$。

解 由于 7/3 不是整数，所以 $3\nmid 7$。另一方面，因为 $12/3=4$，于是 $3\mid 12$。 ■

例 2 令 n 和 d 为正整数，不超过 n 的正整数中有多少个能被 d 整除？

Extra Examples **解** 能被 d 整除的正整数都是 dk 形的，其中 k 是正整数。因此不超过 n 的正整数中能被 d 整除的整数的个数等于能使 $0<dk\leqslant n$ 或 $0<k\leqslant n/d$ 的整数 k 的个数。因此，有 $\lfloor n/d\rfloor$ 个正整数，既不超过 n，又能被 d 整除。 ■

定理 1 给出了整数整除性的某些基本性质。

定理 1 令 a，b，c 为整数。则

1. 若 $a\mid b$ 和 $a\mid c$，则 $a\mid(b+c)$；
2. 若 $a\mid b$，那么对所有整数 c 都有 $a\mid bc$；
3. 若 $a\mid b$，$b\mid c$，则 $a\mid c$。

证 假定 $a\mid b$ 和 $a\mid c$。从整除性定义知有整数 s 和 t，使 $b=as$ 和 $c=at$。因此

$$b+c=as+at=a(s+t)$$

于是 a 整除 $b+c$。这就证明了定理中的(1)。(2)和(3)的证明留给读者做练习。 □

定理 1 有一个有用的推论。

推论 1 如果 a，b，c 是整数使得 $a\mid b$ 和 $a\mid c$，那么只要 m 和 n 是整数，就有 $a\mid mb+nc$。

证 根据定理 1 中的(2)，只要 m 和 n 是整数，就有 $a\mid mb$ 和 $a\mid nc$。根据定理 1 中的(1)，有 $a\mid mb+nc$。 □

2.4.3 素数

大于 1 的每个正整数都至少能被两个整数整除，因为正整数可以被 1 和它自己整除。恰有两个不同的正整数因子的整数称为素数。

定义 2 大于 1 的正整数 p 称为*素数*，如果 p 仅有的正因子是 1 和 p。大于 1 又不是素数的正整数称为*合数*。

注意 整数 n 是合数，当且仅当存在整数 a 使得 $a\mid n$ 并且 $1<a<n$。

例 3 由于 1 和 7 是 7 仅有的正因子，7 是素数。9 是合数，因为它能被 3 整除。 ■

小于 100 的素数包括 2，3，5，7，11，13，17，19，23，29，31，37，41，43，47，53，59，61，67，71，73，79，83，89 和 97。在 6.6 节将介绍一个过程，称为*埃拉托色尼筛法*，可用来求不超过整数 n 的所有素数。

正如算术基本定理所示，素数是构造正整数的积木。定理的证明将在 3.3 节给出。

定理 2（*算术基本定理*） 每个大于 1 的正整数都可以唯一地写为两个或多个素数的乘积，其中素数因子以非递减序出现。

例 4 给出了几个整数的素因子分解。

例4 下面给出100，641，999和1024的素因子分解。

Extra Examples

$100=2\cdot 2\cdot 5\cdot 5=2^2 5^2$

$641=641$

$999=3\cdot 3\cdot 3\cdot 37=3^3\cdot 37$

$1024=2\cdot 2\cdot 2\cdot 2\cdot 2\cdot 2\cdot 2\cdot 2\cdot 2\cdot 2=2^{10}$ ■

证明一个给定的整数是素数往往很重要。例如在密码学中为信息加密的某些方法用到大素数。根据下面的定理可以得到证明整数为素数的一个算法。

定理3 如果 n 是个合数，那么 n 必有小于或等于 $\sqrt{n}$ 的一个素因子。

证 如果 n 是合数，它有一个因子 a，使 $1<a<n$，于是 $n=ab$，其中 a 和 b 都是大于1的正整数，这样就有 $a\leqslant\sqrt{n}$ 或 $b\leqslant\sqrt{n}$，因为否则 $ab>\sqrt{n}\cdot\sqrt{n}=n$。所以 n 有个不大于 $\sqrt{n}$ 的正因子。这个因子或是素数，或根据算术基本定理有素因子。无论哪种情况，n 有个小于或等于 $\sqrt{n}$ 的素因子。□

从定理3可知，如果整数不能被小于或等于其平方根的素数整除，它就是素数。下面的例子就用这一结论证明101是素数。

例5 证明101是素数。

解 不超过 $\sqrt{101}$ 的仅有的素数是2，3，5和7。因为101不能被2，3，5和7整除(101被这些数除的商都不是整数)，所以101是素数。 ■

由于每个整数都有素因子分解。若能找到求解这一素因子分解的算法，一定会很有用。考虑求解整数 n 的素因子分解的问题。从最小素数2开始，从小到大用一个个素数去除 n。如果 n 有素因子，那么必能找到一个不超过 $\sqrt{n}$ 的素因子 p。所以如果找不到不超过 $\sqrt{n}$ 的素因子，n 为素数。否则，如果找到素因子 p，可以对 n/p 继续这样做。注意 n/p 没有小于 p 的素因子。同样的道理，如果 n/p 没有大于等于 p 且不超过它的平方根的素因子，它必为素数。否则如果它有个素因子 q，可以继续对 $n/(pq)$ 做同样的工作。这一过程继续下去直到素数分解就只含一个素数。例6解释了这一过程。

例6 求7007的素因子分解。

解 要求7007的素因子分解，从2开始用一个个素数除7007。2，3和5除不尽7007。但7除尽7007，$7007/7=1001$。下一步从7开始用一个个素数除1001。我们立刻发现7还能除尽1001，$1001/7=143$。从7开始，再用一个个素数除143。虽然7不能整除143，但11可以，$143/11=13$。由于13为素数。这一过程即告完成。由此得7007的素因子分解是 $7\cdot 7\cdot 11\cdot 13=7^2\cdot 11\cdot 13$。 ■

古代因哲学的原因而研究素数。今天研究素数的理由就很实际了。特别是在密码学中，大素数起着关键性的作用。在2.6节将会具体讨论这种作用。

素数的无限性 人们长期以来就已经知道有无限多个素数。我们将用欧几里得在著名的数学教科书《几何原本》中给出的证明来证明这个事实。

定理4 有无限多个素数。

证 用反证法证明这个定理。假设只有有限多个素数 $p_1, p_2, \cdots, p_n$。令

$$Q=p_1p_2\cdots p_n+1$$

根据算术基本定理，Q 要么是素数，要么能被写成两个或多个素数之积。但是素数 p_j 都不能整除 Q，因为如果 $p_j\mid Q$，则 p_j 整除 $Q-p_1p_2\cdots p_n=1$。这是个矛盾，起因是假设列出了所有的素数。于是有无限多个素数。(注意在这个证明中不说 Q 是素数!) □

由于有无限多个素数，所以给定任意正整数，都有大于这个整数的素数。人们一直希望发现越来越大的素数；差不多近300年来已知最大的素数都是形如 2^p-1 的这种特殊的整数，其中 p 也是素数。这种素数称为*梅森素数*，这是以法国修道士梅森[⊖]的名字命名的。梅森在17世纪对这些素数进行了研究。之所以最大已知素数通常都是梅森素数，是因为有个特别有效的称为卢卡斯-莱默尔(Lucas-Lehmer)检验的方法可以判断 2^p-1 是否为素数。而且当前还不可能以差不多同样的速度判断一个不是特殊形式的整数是否为素数。

例7 整数 $2^2-1=3$，$2^3-1=7$ 和 $2^5-1=31$ 都是梅森素数，而 $2^{11}-1=2047$ 不是梅森素数，因为 $2047=23\cdot89$。■

自从发明了计算机，发现梅森素数的工作一直稳步发展。到2002年中期，已知39个不同的梅森素数，其中8个是1990年以来找到的。最大的已知梅森素数(至2002年中期)是 $2^{13\,466\,917}-1$，这是个超过400万位的数。作为一种集体的努力，一个称为"互联网梅森素数大搜索(GIMPS)"的组织已经成立，以共同寻找新的梅森素数。顺便说一句，对梅森素数的寻找这件事本身是有实际意义的。对超级计算机的一种质量控制检验就是重复确定一个大梅森素数的卢卡斯-勒梅检验。

素数的分布 定理4说有无限多个素数。但是有多少个素数小于一个正整数 x 呢？这个问题吸引数学家很多年。在18世纪晚期数学家编制了很大的素数表来收集素数分布的证据。利用这些证据，当时的大数学家包括高斯和勒让德，都猜想有定理5，但是并没有证明。

定理5(*素数定理*) 当 x 无限增长时，不超过 x 的素数个数与 $x/\ln x$ 之比趋于1。(这里 $\ln x$ 是 x 的自然对数。)

法国数学家阿达马(Jacques Hadamard)和比利时数学家瓦列-普金(Charles-Jean-Gustave-Nicholas de la Valleé-Poussin)利用复变函数论在1896年首次证明了素数定理。虽然已经找到了不用复变函数论的证明，但是素数定理的所有已知证明都非常复杂。

可以用素数定理估计一个随机选择的特定规模的数是素数的概率。素数定理说不超过 x 的素数个数可以用 $x/\ln x$ 来逼近。于是一个随机选择的正整数 x 是素数的概率大约是 $(x/\ln x)/x=1/\ln x$。例如，一个靠近 10^{1000} 的整数是素数的概率大约是 $1/\ln 10^{1000}$，即大约1/2300。(当然了，只选择奇数可以使找到素数的概率加倍。)

由定理3知可用简单的除法写出一个因数分解和素数检验的过程。不过这样的过程不是有效的算法；人们已经开发了许多有效得多的算法完成这些任务。因数分解和素数检验对数论在密码学中的应用已变得很重要。这引起了人们极大的兴趣来开发完成这两个任务的有效算法。在过去的30年中已经研究设计了一些聪明的过程，能有效地生成大素数。然而

⊖ **梅森(Marin Mersenne，1588—1648)**生于法国缅因的一个劳动家庭，上过Mans学院和位于La Fle'che的基督学院。后来，从1609年到1611年在Sorbonne继续接受教育，学习神学。1611年他加入了称为Minims的宗教团体。Minims的名字来自Minimi(意为极小——译者)一词(这些人自认为是宗教味最少的团体)。除祈祷外，他们献身于学术和研究。1612年他在巴黎的Place Royale当了牧师。1614年至1618年之间在Nevers的Minims女修道院教哲学。1619年回到巴黎。随后他位于de I'Annociade的Minims宿舍成了包括费马和帕斯卡在内的法国科学家、哲学家和数学家的聚会场所。梅森与全欧洲的学者频繁通信，澄清数学和科学方面的知识，起着数学学术杂志(及今日之互联网)的作用。梅森写的书内容广泛，包括力学、数学物理、数学、音乐和声学。他研究素数，并尝试构造一个能表示所有素数的公式，但没有成功。1644年梅森声称，对 $p=2,3,5,7,13,17,19,31,67,127,257$，$2^p-1$ 是素数，而对于小于257的所有其他素数 2^p-1 不是素数。人们花了300多年的时间在梅森的上述论断中找到5个错误。具体来说，对 $p=67$ 和 $p=257$，2^p-1，不是素数，而对 $p=61$，$p=87$ 和 $p=107$，2^p-1 是素数。还值得一提的是，梅森为受到宗教批评的两位当时最出名的人辩护，这两个人就是笛卡儿和伽利略。他还协助揭露炼丹术士和占星术士的骗术。

在同一时期中，尽管已开发了有力的因数分解新方法，大整数分解因数仍然要花费特别长的时间。不过分解大整数为素数的挑战引起了许多人的兴趣。互联网上有集体的努力来分解大整数，特别是形如 $k^n \pm 1$ 的大数，其中 k 是个小的正整数而 n 是个大正整数(这种数称为卡宁汉数)。任何时候，总有一组称为“最热十个”的这种大整数等待分解。

2.4.4 整除算法

当一个整数除以一个正整数时，有一个商和一个余数，整除算法说明这一点。

定理 6(整除算法) 令 a 为整数，d 为正整数，那么有唯一的整数 q 和 r，其中 $0 \leqslant r < d$，使得 $a = dq + r$。

注意 定理 6 实际上并不是一个算法。(为什么不是?)我们采用的是它的传统称呼。

定义 3 在上述整除算法给出的等式中，d 称为除数，a 称为被除数，q 称为商，r 称为余数。用下面这种记号表示商和余数：

$$q = a \textbf{ div } d, \quad r = a \textbf{ mod } d。$$

例 8 和例 9 解释了整除算法。

例 8 101 除以 11 时商和余数是什么?

解 我们有

$$101 = 11 \cdot 9 + 2$$

所以 101 除以 11 时商是 $9 = 101$ **div** 11，余数是 $2 = 101$ **mod** 11。 ■

例 9 -11 除以 3 时商和余数是什么?

Extra Examples

解 我们有

$$-11 = 3 \cdot (-4) + 1。$$

所以 -11 除以 3 时商是 $-4 = -11$ **div** 3，余数是 $1 = -11$ **mod** 3。

注意余数不能为负。因此余数不是 -2，即使

$$-11 = 3 \cdot (-3) - 2,$$

因为 $r = -2$ 不满足 $0 \leqslant r < 3$。 ■

注意整数 a 能被整数 d 整除的充分必要条件是当 a 除以 d 时余数为 0。

2.4.5 最大公约数和最小公倍数

能整除两个整数的最大整数称为这两个整数的最大公约数。

定义 4 令 a 和 b 是不全为 0 的两个整数。能使 $d \mid a$ 和 $d \mid b$ 的最大整数 d 称为 a 和 b 的最大公约数。a 和 b 的最大公约数用 $\gcd(a, b)$ 表示。

不全为 0 的两个整数的最大公约数的确存在，因为这两个整数的公约数集合是有限的，求两个整数的最大公约数的一个方法是求出两个整数的所有正的公约数，然后取其中最大的。下面的例子就是这样做的，随后会给出一个更有效的求最大公约数的方法。

例 10 24 和 36 的最大公约数是什么?

解 24 和 36 的正公约数是 1，2，3，4，6 和 12，所以 $\gcd(24, 36) = 12$。 ■

例 11 17 和 22 的最大公约数是什么?

解 17 和 22 除 1 以外没有正公约数，所以 $\gcd(17, 22) = 1$。 ■

由于往往需要指明两个整数除 1 以外没有其他正公约数，所以给出下面的定义。

定义 5 如果整数 a 和 b 的最大公约数是 1，就说它们是互素的。

例 12 从例 11 可知整数 17 和 22 是互素的，因为 gcd(17，22)＝1。■

由于经常需要指明一个整数集合中任何两个整数都没有大于 1 的正公约数，所以给出下面的定义。

定义 6 整数 a_1，a_2，…，a_n 是两两互素的，如果只要 $1\leqslant i<j\leqslant n$，就有 $\gcd(a_i, a_j)=1$。

例 13 判断整数 10，17 和 21 是否两两互素，整数 10，19 和 24 是否两两互素。

解 由于 gcd(10，17)＝1，gcd(10，21)＝1 和 gcd(17，21)＝1，所以断定 10，17 和 21 是两两互素的。

由于 gcd(10，24)＝2＞1，所以 10，19 和 24 不是两两互素的。■

求两个整数的最大公约数的另一方法是利用这两个整数的素因子分解。假定两个全不为 0 的整数 a 和 b 的素因子分解为：

$$a = p_1^{a_1} p_2^{a_2} \cdots p_n^{a_n}, \quad b = p_1^{b_1} p_2^{b_2} \cdots p_n^{b_n}$$

其中每个指数都是非负整数，而且出现在 a 和 b 分解中的所有素数都包含在两个分解之中，必要时以 0 为指数出现。于是 $\gcd(a, b)$ 由下面的公式给出：

$$\gcd(a,b) = p_1^{\min(a_1,b_1)} p_2^{\min(a_2,b_2)} \cdots p_n^{\min(a_n,b_n)}$$

其中 $\min(x, y)$ 代表两个数 x 和 y 的最小值。为证明这一计算 $\gcd(a, b)$ 的公式是有效的，必须证明等式右边的整数整除 a 和 b，而且没有比它大的数也能整除 a 和 b。由于右边这个数中每个素数的指数都不超过 a 和 b 的分解中该素数的指数，所以它的确整除 a 和 b。进一步说，没有更大的整数解整除 a 和 b，因为分解式中每个素数的指数都不能再增大，而且其他素数也不能加进来。

例 14 由于 120 和 500 的素因子分解分别是 $120=2^3\cdot 3\cdot 5$ 和 $500=2^2\cdot 5^3$，所以它们的最大公约数是

$$\gcd(120,500) = 2^{\min(3,2)}\cdot 3^{\min(1,0)}\cdot 5^{\min(1,3)} = 2^2\cdot 3^0\cdot 5^1 = 20$$ ■

因数分解亦可用于求两个整数的最小公倍数。

定义 7 正整数 a 和 b 的最小公倍数是能被 a 和 b 整除的最小正整数。a 和 b 的最小公倍数用 $\operatorname{lcm}(a, b)$ 表示。

最小公倍数的存在是由于能被 a 和 b 整除的整数之集合是非空的，而每个非空的正整数集合都有一个最小元素(根据 3.3 节将讨论的良序性质)。假定 a 和 b 的素因子分解如前面给出的那样，那么 a 和 b 的最小公倍数由下式给出：

$$\operatorname{lcm}(a,b) = p_1^{\max(a_1,b_1)} p_2^{\max(a_2,b_2)} \cdots p_n^{\max(a_n,b_n)}$$

其中 $\max(x, y)$ 表示两个数 x 和 y 中的最大数。这一公式是有效的，因为 a 和 b 的公倍数在其分解中至少含 $\max(a_i, b_i)$ 个 p_i，而最小公倍数的分解中又没有 a 和 b 的因子之外的素数。

例 15 $2^3 3^5 7^2$ 和 $2^4 3^3$ 的最公倍数是什么?

解 我们有

$$\operatorname{lcm}(2^3 3^5 7^2, 2^4 3^3) = 2^{\max(2,4)} 3^{\max(5,3)} 7^{\max(2,0)} = 2^4 3^5 7^2$$ ■

下面的定理给出了两个整数的最大公约数和最小公倍数之间的关系。用上面给出的求这两个数的公式就可以证明这一定理。具体证明留给读者作为练习。

定理 7 令 a 和 b 为正整数，则

$$ab = \gcd(a,b)\cdot \operatorname{lcm}(a,b)$$

2.4.6 同余算术

有时我们只关心一个整数除以另一个指定的正整数时的余数。例如，如果问从现在开始 50 个小时以后的时间(每天 24 小时计时)，那么指的是当前钟点数加上 50 除以 24 的余数。由于往

往只对余数感兴趣，所以用专门的记号表示余数。

我们用一个记号表示两个整数除以正整数 m 时有同样的余数。

定义 8 若 a 和 b 为整数而 m 为正整数，如果 m 整除 $a-b$，就说 a 模 m 同余 b。用 $a\equiv b(\mathbf{mod}\ m)$表示 a 模 m 同余 b。如果 a 和 b 不是模 m 同余的，就写成 $a\not\equiv b(\mathbf{mod}\ m)$。

定理 8 说明了处理余数时所用的概念之间的联系。

定理 8 令 a 和 b 为整数而 m 为正整数。则 $a\equiv b(\mathbf{mod}\ m)$当且仅当 $a\ \mathbf{mod}\ m=b\ \mathbf{mod}\ m$。

定理 8 的证明留做本节末尾的练习 21 和 22。

例 16 判断 17 是否模 6 同余于 5，24 是否模 6 同余于 14。

解 由于 6 整除 $17-5=12$，所以 $17\equiv 5(\mathbf{mod}\ 6)$。但 $24-14=10$ 不能被 6 整除，所以 $24\not\equiv 14(\mathbf{mod}\ 6)$。■

伟大的德国数学家高斯[⊖]在 18 世纪末发展了同余的概念。

同余概念在数论发展中起着重要的作用。下面的定理给出了一个使用同余的方法。

定理 9 令 m 为正整数，整数 a 和 b 模 m 同余的充分必要条件是存在整数 k，使 $a=b+km$。

证 若 $a\equiv b(\mathbf{mod}\ m)$，那么 $m\mid(a-b)$。这表示有整数 k 使 $a-b=km$，于是 $a=b+km$。反过来，若有整数 k 使 $a=b+km$，那么 $km=a-b$。于是 m 整除 $a-b$，所以 $a\equiv b(\mathbf{mod}\ m)$。□

整数 a 模 m 所有整数同余的集合称为 a 模 m 的同余类。在第 7 章将证明有 m 对模 m 的不相交等价类，这些等价类的并是整数的集合。

定理 10 说明同余和加法乘法的关系。

定理 10 令 m 为正整数。若 $a\equiv b(\mathbf{mod}\ m)$，$c\equiv d(\mathbf{mod}\ m)$，那么

$$a+c\equiv b+d(\mathbf{mod}\ m)\quad 及\quad ac\equiv bd(\mathbf{mod}\ m)$$

证 因为 $a\equiv b(\mathbf{mod}\ m)$，$c\equiv d(\mathbf{mod}\ m)$，所以有整数 s 和 t 使 $b=a+sm$ 和 $d=c+tm$。于是

$$b+d=(a+sm)+(c+tm)=(a+c)+m(s+t)$$

及

$$bd=(a+sm)(c+tm)=ac+m(at+cs+stm)$$

因此

$$a+c\equiv b+d(\mathbf{mod}\ m)\quad 及\quad ac\equiv bd(\mathbf{mod}\ m)$$ □

例 17 由于 $7\equiv 2(\mathbf{mod}\ 5)$和 $11\equiv 1(\mathbf{mod}\ 5)$，从定理 10 知

$$18=7+11\equiv 2+1=3(\mathbf{mod}\ 5)$$

且

$$77=7\cdot 11\equiv 2\cdot 1=2(\mathbf{mod}\ 5)$$ ■

2.4.7 同余应用

数论有广泛的应用，本节将介绍其中三个：用同余为计算机文件分配内存地址、生成伪随机数，以及基于同余算术的密码系统。



⊖ **高斯(Karl Friedrich Gauss，1777—1855)** 泥瓦匠的儿子高斯是个神童，10 岁时即展现了非凡的潜力，迅速解答了老师出的难题。当时老师要学生计算头一百个正整数的和。高斯发现这一百个数可以分成五十对，每一对都是 101：100+1，99+2，…，50+51。智慧的火花引起了包括布朗斯威克公爵费迪南德在内的要人的注意。在他们的提议下，高斯得以在卡洛林学院和哥廷根大学学习。还是个学生时，他就发明了从实验结果估计变量最可能的值的最小平方法。1796 年高斯做出了几何学上的基本发现，推动了自古代已停步不前的这一学科的发展。他证明了只用圆规和直尺可以画正 17 边形。

1799 年高斯给出了算术基本定理第一个严格的证明。这一定理指出 n 次多项式恰有 n 个根(重根以重数计算)。当他成功地用不充分的数据计算出人类首次发现的小行星谷神星的轨道时，高斯赢得了世界声誉。

高斯被他同时代的数学家称为数学王子。尽管高斯以其在几何、代数、数学分析、天文和物理上的许多发现而知名，他对数论有着特别的兴趣。这从他的名言可见一斑："数学是科学的皇后，而数论是数学的皇后。"1801 年高斯出版了《*Disquisitiones Arithmeticae*》一书，为现代数论奠定了基础。

例 18(散列函数) 学校的中央计算机保存着所有学生的档案记录。怎样分配内存地址才能迅速检索到学生记录？适当选一个*散列函数*就是这一问题的解。记录用关键字识别，每个关键字唯一地识别一个学生记录。例如往往用学生的社会安全号作为他的记录的关键字。散列函数 h 将内存地址 $h(k)$ 分配给以 k 为关键字的记录。

实践中使用许多不同的散列函数。最常用的散列函数之一是

$$h(k) = k \ \mathbf{mod}\ m$$

其中 m 是可供使用的内存地址的数目。

散列函数应该易于计算以便快速检索到文件。散列函数 $h(k)=k \ \mathbf{mod}\ m$ 符合这一要求：要求 $h(k)$，只须计算 k 被 m 除的余数。散列函数还应该是满射的，这样所有内存地址均可利用。函数 $h(k)=k \ \mathbf{mod}\ m$ 也符合这一要求。

例如，当 $m=111$ 时，以 064212848 为社会安全号的学生的记录分配到的地址是 14，因为

$$h(064212848) = 064212848 \ \mathbf{mod}\ 111 = 14$$

类似地，由于

$$h(037149212) = 037149212 \ \mathbf{mod}\ 111 = 65$$

以 037149212 为社会安全号的学生的记录得到的内存地址是 65。

由于散列函数不是一对一的(因为关键字的个数可能大于内存地址数)，有可能多个记录分配到同一个内存地址。这时就说出现了*冲突*。消解冲突的一个办法是使用散列函数给出的但已被占用的地址后面第一个未占用地址。例如在分配了上述两个地址以后，把 15 分配给社会安全号为 107405723 的学生记录。要看清这一点，首先注意 $h(k)$ 会把这一社会安全号映射到地址 14，这是因为

$$h(107405723) = 107405723 \ \mathbf{mod}\ 111 = 14$$

但这一地址已被占用(社会安全号为 064212848 的学生文件占用)，而 15 是内存地址 14 后面第一个未占用的地址。

有许多较上述简单方法复杂的消解冲突的办法，本书结尾处给出的有关散列函数的参考文献讨论了这些方法。 ■

例 19(伪随机数) 计算机模拟常需要用到随机选择的数。有不同的方法可以产生具有随机中选性质的数。由于用系统的方法产生的数不可能真正是随机的，就称为*伪随机数*。

最常用的产生伪随机数的过程称为*线性同余法*。选择 4 个数：*模数* m，*乘数* a，*增量* c 和*种子* x_0，使得 $2 \leqslant a < m$，$0 \leqslant c < m$ 及 $0 \leqslant x_0 < m$。生成一个伪随机数序列 $\{x_n\}$ 使得对所有 n，$0 \leqslant x_n < m$。生成的办法是逐次同余：

$$x_{n+1} = (ax_n + c) \ \mathbf{mod}\ m$$

(这是个递归定义的例子，递归定义将在 3.4 节讨论。在这一节只证明这样定义的序列是定义良好的。)

不少计算机试验都要求产生 0 和 1 之间的伪随机数。要得到这样的数，可以用线性同余法生成的数除以模数，即使用 x_n/m。

例如选 $m=9$，$a=7$，$c=4$ 和 $x_0=3$，产生的伪随机数序列如下：

$$x_1 = 7x_0 + 4 = 7.3 + 4 = 25 \ \mathbf{mod}\ 9 = 7$$
$$x_2 = 7x_1 + 4 = 7.7 + 4 = 53 \ \mathbf{mod}\ 9 = 8$$
$$x_3 = 7x_2 + 4 = 7.8 + 4 = 60 \ \mathbf{mod}\ 9 = 6$$
$$x_4 = 7x_3 + 4 = 7.6 + 4 = 46 \ \mathbf{mod}\ 9 = 1$$

$$x_5 = 7x_4 + 4 = 7.1 + 4 = 11 \ \mathbf{mod}\ 9 = 2$$
$$x_6 = 7x_5 + 4 = 7.2 + 4 = 18 \ \mathbf{mod}\ 9 = 0$$
$$x_7 = 7x_6 + 4 = 7.0 + 4 = 4 \ \mathbf{mod}\ 9 = 4$$
$$x_8 = 7x_7 + 4 = 7.4 + 4 = 32 \ \mathbf{mod}\ 9 = 5$$
$$x_9 = 7x_8 + 4 = 7.5 + 4 = 39 \ \mathbf{mod}\ 9 = 3$$

由于 $x_9 = x_0$ 而且每一项都只依赖于其前面一项，所以产生的序列如下：

$$3,7,8,6,1,2,0,4,5,3,7,8,6,1,2,0,4,5,3,\cdots$$

这个序列含 9 个不同的数，然后重复。

大部分计算机的确使用线性同余法生成伪随机数。常使用的线性同余发生器的增量 $c=0$。这样的发生器称为*纯乘式发生器*。例如以 $2^{31}-1$ 为模，以 $7^5=16\,807$ 为乘数的纯乘式发生器就广为采用。可以证明以这些值来计算，会产生 $2^{31}-2$ 个数，然后开始重复。 ■

2.4.8 密码学

同余在离散数学和计算机科学中有许多应用。可以在本书末尾给出的推荐读物中找到对这些应用的讨论。最重要的同余应用之一涉及研究信息保密的*密码学*。已知最早使用密码的例子之一是凯撒(Julius Caesar)。他把字母表中的每个字母往前移动三位以获得保密信息(字母表的头三个字母移为最后三个)。例如根据这一方案，字母 E 移到 B，而字母 A 移到 X。这是*加密*的一个例子，加密就是产生保密信息的过程。

要用数学来表达凯撒的加密过程，就按照字母在字母表中的位置用 0 到 25 的数表示字母。例如用 0 表示 A，用 10 表示 K，用 25 表示 Z。凯撒的加密方法可以用函数 f 表示，对每个非负整数 p，$p \leqslant 25$，函数值 $f(p)$ 是 $\{0, 1, 2, \cdots, 25\}$ 中的一个数，使

$$f(p) = (p+3) \ \mathbf{mod}\ 26$$

在加密信息中，由 p 代表的字母用 $(p+3)$ **mod** 26 代表的字母代替。

例 20 从信息“MEET YOU IN THE PARK”用凯撒密码获得的加密信息是什么？

解 首先用数代替信息中的字母，得：

12 4 4 19　　24 14 20　　8 13　　19 7 4　　15 0 17 10

再用 $f(p)=(p+3)$ **mod** 26 代替 p，得：

15 7 7 22　　1 17 23　　11 16　　22 10 7　　18 3 20 13

翻译成字母即得加密信息“PHHW BRX LQ WKH SDUN”。 ■

要把凯撒密码加密的保密信息还原，需使用 f 的反函数 f^{-1}。注意 f^{-1} 把 $\{0, 1, 2, \cdots, 25\}$ 中的整数 p 变成 $f^{-1}(p)=(p-3)$ **mod** 26。换句话说，要找回原信息，每个字母要在字母表中向后移三位，最后三个字母移到前三位。从加密信息恢复成原信息的过程称为*解密*。

有各种方法可以扩展凯撒密码。例如不把每个字母移动三位，而是移动 k 位，于是

$$f(p) = (p+k) \ \mathbf{mod}\ 26$$

这样的密码称为*移位密码*。注意解密可以用

$$f^{-1}(p) = (p-k) \ \mathbf{mod}\ 26$$

来完成。

显然，凯撒的方法和移位密码不能提供高度安全。有各种改进这一方法的办法。稍稍提高一点安全性的一种办法是使用形为

$$f(p) = (ap+b) \ \mathbf{mod}\ 26$$

的函数，其中 a 和 b 为整数，而且需保证 f 为双射。(这样的映射称为*仿射变换*。)这种函数产生

若干可能的加密系统。下面的例子解释了如何使用一个这样的系统。

例 21 用 $f(p)=(7p+3)$ **mod** 26 加密时，用什么字母取代字母 K？

解 首先注意 10 代表 K。然后用指定的加密函数得到 $f(10)=(7\cdot 10+3)$ **mod** $26=21$。因为 21 代表 V，在加密信息中用 V 代表字母 K。■

凯撒的加密法及其扩展都是用字母表中的一个字母代替另一个。这类加密法很容易受到根据字母在信息中出现的频率作出的攻击。复杂些的加密法是用一段字母取代另一段字母。已经有许多种根据同余算术设计的加密字母段的技术。对这些技术的讨论可以在本书末列出的推荐读物中找到。

练习

1. 17 能整除下列各数吗？
   a)68　　b)84　　c)357　　d)1001
2. 如果 a 是不为 0 的整数，证明
   a)1 整除 a。　　b)a 整除 0。
3. 证明定理 1 的第 2 部分成立。
4. 证明定理 1 的第 3 部分成立。
5. 证明如果 $a\mid b$ 及 $b\mid a$，其中 a 和 b 为整数，则必有 $a=b$ 或 $a=-b$。
6. 证明若 a、b、c 和 d 为整数，使 $a\mid c$ 及 $b\mid d$，则 $ab\mid cd$。
7. 若 a，b，c 为整数，使 $ac\mid bc$，求证 $a\mid b$。
8. 下列整数是素数吗？
   a)19　　b)27　　c)93　　d)101　　e)107　　f)113
9. 求下列各种情况的商和余数。
   a)19 除以 7　　b)−111 除以 11　　c)789 除以 23　　d)1001 除以 13
   e)0 除以 19　　f)3 除以 5　　g)−1 除以 3　　h)4 除以 1
10. 求下列各种情况的商和余数。
   a)44 除以 8　　b)777 除以 21　　c)−123 除以 19　　d)−1 除以 23
   e)−2002 除以 87　　f)0 除以 17　　g)1 234 567 除以 1001　　h)−100 除以 101
11. 求下列各数的素因数分解。
   a)88　　b)126　　c)729　　d)1001　　e)1111　　f)909 090
12. 求下列各数的素因数分解。
   a)39　　b)81　　c)101　　d)143　　e)289　　f)899
13. 求 10! 的素数分解。

*14. 100! 的尾部有多少个 0？

*15. 求证 $\log_2 3$ 是无理数。回忆一下，无理数是不能写成两个整数之比的实数。

16. 小于 12 的哪些正整数与 12 互素？
17. 小于 30 的哪些正整数与 30 互素？
18. 判断下列各组整数是否两两互素？
   a)21，34，55　　b)14，17，85　　c)25，41，49，64　　d)17，18，19，23
19. 判断下列各组整数是否两两互素？
   a)11，15，19　　b)14，15，21　　c)12，17，31，37　　d)7，8，9，11
20. 如果一个正整数等于除它自己以外它的所有正因子的和，这个数就称为完美的。
   a)证明 6 和 28 是完美的。
   b)当 2^p-1 为素数时，求证 $2^{p-1}(2^p-1)$ 是完美的。

21. 令 m 为正整数。求证若 a **mod** $m=b$ **mod** m，则 $a\equiv b(\textbf{mod}\ m)$。
22. 令 m 为正整数。求证若 $a\equiv b(\textbf{mod}\ m)$，则 a **mod** $m=b$ **mod** m。
23. 求证若 2^n-1 为素数，则 n 为素数。（提示：利用等式 $2^{ab}-1=(2^a-1)2^{a(b-1)}+(2^{a(b-2)}+\cdots+2^a+1)$。）
24. 判断下列各整数是否为素数，以此证明梅森的某些论断。

    a) 2^7-1　b) 2^9-1　c) $2^{11}-1$　d) $2^{13}-1$
25. 欧拉 ϕ 函数在正整数 n 的值定义为小于或等于 n 且与 n 互素的正整数的个数。（注意：ϕ 是希腊字母。）求

    a) $\phi(4)$　b) $\phi(10)$　c) $\phi(13)$
26. 求证 n 为素数的充分必要条件是 $\phi(n)=n-1$。
27. $\phi(p^k)$ 的值是什么，其中 p 为素数，k 为正整数？
28. 下列各对整数的最大公约数是什么？

    a) $2^2\cdot 3^3\cdot 5^5$，$2^5\cdot 3^3\cdot 5^2$　b) $2\cdot 3\cdot 5\cdot 7\cdot 11\cdot 13$，$2^{11}\cdot 3^9\cdot 11\cdot 17^{14}$

    c) 17，17^{17}　d) $2^2\cdot 7$，$5^3\cdot 13$

    e) 0，5　f) $2\cdot 3\cdot 5\cdot 7$，$2\cdot 3\cdot 5\cdot 7$
29. 下列各对整数的最大公约数是什么？

    a) $3^7\cdot 5^3\cdot 7^3$，$2^{11}\cdot 3^5\cdot 5^9$　b) $11\cdot 13\cdot 17$，$2^9\cdot 3^7\cdot 5^5\cdot 7^3$

    c) 23^{31}，23^{17}　d) $41\cdot 43\cdot 53$，$41\cdot 43\cdot 53$

    e) $3^{13}\cdot 5^{17}$，$2^{12}\cdot 7^{21}$　f) 1111，0
30. 练习 28 中各对整数的最小公倍数是什么？
31. 练习 29 中各对整数的最小公倍数是什么？
32. 试求 gcd(1000，625)和 lcm(1000，625)，并验证 gcd(1000，625)・lcm(1000，625)＝1000・625。

*33. 若 n 和 k 为正整数，求证 $\lceil n/k\rceil=\lfloor (n-1)/k\rfloor+1$。

34. 若 a 为整数，d 是大于 1 的正整数，求证 a 除以 d 的商和余数分别是 $\lfloor a/d\rfloor$ 和 $a-d\lfloor a/d\rfloor$。
35. 设 m 为正整数，给出计算与整数 a 模 m 同余的绝对值最小的整数的公式。
36. 计算下列各量：

    a) -17 **mod** 2　b) 144 **mod** 7　c) -101 **mod** 13　d) 199 **mod** 19
37. 计算下列各量：

    a) 13 **mod** 3　b) -97 **mod** 11　c) 155 **mod** 19　d) -221 **mod** 23
38. 列出 5 个与 4 模 12 同余的整数。
39. 判断下列各整数是否模 17 同余于 5。

    a) 80　b) 103　c) -29　d) -122
40. 若两个整数的乘积为 $2^7 3^8 5^2 7^{11}$，它们的最大公约数为 $2^3 3^4 5$，它们的最小公倍数是什么？
41. 若 a 和 b 为正整数，求证 $ab=\gcd(a,b)\cdot \text{lcm}(a,b)$。（提示：利用 a 和 b 的素数分解以及根据素数分解给出的计算 $\gcd(a,b)$ 和 $\text{lcm}(a,b)$ 的公式。）
42. 若 $a\equiv b(\text{mod}\ m)$，$c\equiv d(\text{mod}\ m)$，其中 a，b，c，d 和 m 为整数，且 $m\geqslant 2$，求证 $a-c\equiv b-d\ (\text{mod}\ m)$。
43. 若 $n\mid m$，m 为大于 1 的正整数，且 $a\equiv b(\text{mod}\ m)$，其中 a，b 为整数，求证 $a\equiv b(\text{mod} n)$。
44. 求证若 a，b，c 和 m 为整数，$m\geqslant 2$，$c>0$ 且 $a\equiv b(\text{mod}\ m)$，则 $ac\equiv bc(\text{mod}\ mc)$。
45. 若 a，b，c 和 m 为整数，$m\geqslant 2$，证明 $ac\equiv bc(\text{mod}\ m)$ 不一定蕴含 $a\equiv b(\text{mod}\ m)$。
46. 若 a，b 和 m 为整数，$m\geqslant 2$ 且 $a\equiv b(\text{mod}\ m)$，求证 $\gcd(a,m)=\gcd(b,m)$。
47. 若 a，b，k，m 为整数，$k\geqslant 1$，$m\geqslant 2$ 且 $a\equiv b(\text{mod}\ m)$，求证 $a^k\equiv b^k(\text{mod}\ m)$。
48. 若用散列函数 $h(k)=k$ **mod** 101 为学生记录分配地址，下列社会安全号得到的地址是什么？

    a) 104578690　b) 432222187　c) 372201919　d) 501338753
49. 停车场有 31 个车位供来访者使用，编号为 0 到 30。来访者根据散列函数 $h(k)=k$ mod 31 获得车位，其中 k 是来访者车牌前三位数。

a)下列车牌前三位数得到什么车位？

317，918，007，100，111，310

b)给出一个办法，使来访者在发现他们得到的车位已被占用时可以找到空车位。

50. 用线性同余发生器 $x_{n+1}=(4x_n+1)$ **mod** 7 和种子 $x_0=3$ 产生的伪随机数序列是什么？
51. 用纯乘式发生器 $x_{n+1}=3x_n$ **mod** 11 和种子 $x_0=2$ 产生的伪随机数序列是什么？
52. 写一伪码算法，用线性同余发生器产生伪随机数序列。
53. 为信息"DO NOT PASS GO"加密，先把字母翻译成整数，再用下面给的加密函数计算，然后把整数翻译成字母。

    a) $f(p)=(p+3)$ **mod** 26 （凯撒密码）

    b) $f(p)=(p+13)$ **mod** 26

    c) $f(p)=(3p+7)$ **mod** 26
54. 为下列用凯撒密码加密的信息解密。

    a)EOXH MHDQV　　b)WHVW WRGDB　　c)HDW GLP VXP

书籍用国际标准书号(ISBN)来识别，这是个十位编码，$x_1x_2\cdots x_{10}$，由出版社给定。这十位码分段表示语言、出版社、出版社给书的编号，以及最后一个校验位。校验位或是数字，或是字母 X(用来表示 10)。校验位用 $\sum_{i=1}^{10} ix_i\equiv 0(\text{mod } 11)$来选择，用于发现各位数字的错误或数字移位。

55. 本书第 3 版的 ISBN 前九位数字是 0-07-053965，它的校验位是什么？
56. 《初等数论及其应用》第 3 版的 ISBN 是 0-201-57Q89-1，其中 Q 是个数字。求 Q 的值。
57. 判断出版者给出的本书 ISBN 的校验位是否计算正确。
58. 求恰有 n 个不同因数的最小正整数，其中 n 是

    a)3　　b)4　　c)5　　d)6　　e)10
59. 给出与素数或素数分解有关的公式或规则，用以计算序列的第 n 项，使序列的初始项为下面给出的这些。

    a)0，1，1，0，1，0，1，0，0，0，1，0，1，…

    b)1，2，3，2，5，2，7，2，3，2，11，2，13，2，…

    c)1，2，2，3，2，4，2，4，3，4，2，6，2，4，…

    d)1，1，1，0，1，1，1，0，0，1，1，0，1，1，…

    e)1，2，3，3，5，5，7，7，7，7，11，11，13，13，…

    f)1，2，6，30，210，2310，30 030，510 510，9 699 690，223 092 870，…
60. 给出与素数或素数分解有关的公式或规则，用以计算序列第 n 项，使序列的初始项为下面给出的这些。

    a)2，2，3，5，5，7，7，11，11，11，11，13，13，…

    b)0，1，2，2，3，3，4，4，4，4，5，5，6，6，…

    c)1，0，0，1，0，1，0，1，1，1，0，1，0，1，…

    d)1，−1，−1，0，−1，1，−1，0，0，1，−1，0，−1，1，1，…

    e)1，1，1，1，1，0，1，1，1，0，1，0，1，0，0，…

    f)4，9，25，49，121，169，289，361，529，841，961，1369，…

2.5 整数和算法

2.5.1 引言

正如 2.1 节所说，算法这一术语最初指的是用整数的十进制表示进行算术运算的过程。修改后能处理二进制表示的这些算法是计算机算术的基础。这些算法为理解算法这一概念及算法复

杂度提供了很好的实例。因此本节将讨论这些算法。

除算术中常用的整数算法以外，还有许多涉及整数的重要算法，包括欧几里得算法，这是最有用的算法之一，很可能还是数学中最古老的算法。我们还将描述一个算法，用于对任意基数 b 求正整数的 b 进制展开和求同余幂，在密码学中这是个重要的算法。

2.5.2 整数表示

日常生活中都用十进制记号表示整数。例如，965 用来表示 $9\cdot 10^2+6\cdot 10+5$。不过有时用 10 以外的数为基数更方便。特别是计算机通常用二进制记号(以 2 为基数)来做算术运算，而用八进制(基数为 8)或十六进制(基数为 16)记号来表示字符，如字母或数字。事实上可以用 1 以外的任何正整数为基数来表示整数。定理 1 陈述的就是这一结论。

定理 1 令 b 为不等于 1 的正整数。那么如果 n 是个正整数，就可以唯一地表示为下面的形式：

$$n=a_kb^k+a_{k-1}b^{k-1}+\cdots+a_1b+a_0$$

其中 k 是非负整数，a_0，a_1，…，a_k 是小于 b 的非负整数，$a_k\neq 0$。

本定理的证明可以在[Ro99]中找到。定理 1 中给出的 n 的表示称为 *n 的 b 进制展开*。n 的 b 进制展开表示为 $(a_ka_{k-1}\cdots a_1a_0)_b$。例如，$(245)_8$ 表示 $2\cdot 8^2+4\cdot 8+5=165$。

二进制展开 选择 2 为基数就给出整数的*二进制展开*。在二进制记号中每位数字是 0 或 1。换言之，整数的二进制展开就是个二进制串。计算机用二进制展开(及作为二进制展开变种的相关展开)表示整数并做整数算术。

例 1 以 $(101011111)_2$ 为二进制展开的整数之十进制展开是什么？

解 我们有

$$(101011111)_2=1\cdot 2^8+0\cdot 2^7+1\cdot 2^6+0\cdot 2^5+1\cdot 2^4+1\cdot 2^3+1\cdot 2^2+1\cdot 2^1+1\cdot 2^0=351$$

■

十六进制展开 16 是计算机科学中使用的另一基数。整数以 16 为基数的展开称为*十六进制展开*。这种展开要用 16 个不同的数字。通常使用的数字是 0，1，2，3，4，5，6，7，8，9，A，B，C，D，E 和 F，其中 A 到 F 表示的数字对应于(十进制的)10 到 15。

例 2 十六进制展开 $(2\text{AE0B})_{16}$ 的十进制展开是什么？

解 我们有

$$(2\text{AE0B})_{16}=2\cdot 16^4+10\cdot 16^3+14\cdot 16^2+0\cdot 16+11=(175627)_{10}$$ ■

每个十六进制数字可以用 4 个二进制位表示。例如，可以看出 $(11100101)_2=(\text{E5})_{16}$，因为 $(1110)_2=(\text{E})_{16}$ 而 $(0101)_2=(5)_{16}$。字节是长度为 8 的二进制串，所以字节可以用两个十六进制数字表示。

进制转换 现在介绍一个构造整数 n 的 b 进制展开的算法。首先，用 b 除 n 得到商和余数，即

$$n=bq_0+a_0,\quad 0\leqslant a_0<b$$

余数 a_0 就是 n 的 b 进制展开的最右边一位数字。下一步用 b 除 q_0 得

$$q_0=bq_1+a_1,\quad 0\leqslant a_1<b$$

可以看出 a_1 是 n 的 b 进制展开中从右边数的第二个数字。继续这一过程，不断用商数除以 b 并以余数为新的 b 进制数字。这一过程在商为 0 时终止。

例 3 求 $(12345)_{10}$ 的八进制展开。

解 首先用 8 除 12345，得

$$12345 = 8 \cdot 1543 + 1$$

不断用 8 除商数，得

$$1543 = 8 \cdot 192 + 7$$
$$192 = 8 \cdot 24 + 0$$
$$24 = 8 \cdot 3 + 0$$
$$3 = 8 \cdot 0 + 3$$

由于这些余数就是 12345 的八进制展开中的数字，于是

$$(12345)_{10} = (30071)_8$$

■

例 4 求$(177130)_{10}$的十六进制展开。

解 首先用 16 除 177130，得

$$177130 = 16 \cdot 11070 + 10$$

不断用 16 除商数，得

$$11070 = 16 \cdot 691 + 14$$
$$691 = 16 \cdot 43 + 3$$
$$43 = 16 \cdot 2 + 11$$
$$2 = 16 \cdot 0 + 2$$

由于这些余数就是$(177130)_{10}$的十六进制(基数 16)展开中的数字，于是

$$(177130)_{10} = (2B3EA)_{16}$$

(回忆一下，整数 10，11 和 14 分别对应于十六进制数字 A，B 和 E。) ■

例 5 求$(241)_{10}$的二进制展开。

解 首先用 2 除 241，得

$$241 = 2 \cdot 120 + 1$$

不断用 2 除商数，得

$$120 = 2 \cdot 60 + 0$$
$$60 = 2 \cdot 30 + 0$$
$$30 = 2 \cdot 15 + 0$$
$$15 = 2 \cdot 7 + 1$$
$$7 = 2 \cdot 3 + 1$$
$$3 = 2 \cdot 1 + 1$$
$$1 = 2 \cdot 0 + 1$$

由于这些余数就是$(241)_{10}$的二进制(基数 2)展开中的数字，于是

$$(241)_{10} = (11110001)_2$$

■

算法 1 中给出的伪码用于计算整数 n 的 b 进制展开$(a_{k-1}\cdots a_1 a_0)_b$。

算法 1 构造 b 进制展开

```
procedure base b expansion (n: 正整数)
q := n
k := 0
while q≠0
begin
    a_k := q mod b
    q := ⌊q/b⌋
    k := k+1
end{n 的 b 进制展开是(a_{k-1}…a_1 a_0)_b}
```

在算法 1 中，q 表示不断用 b 去除时得到的商，初值 $q=n$。b 进制展开中的数字就是做这些除法时得到的余数，由 q **mod** b 给出。在得到的商 $q=0$ 时，算法结束。

注意 算法 1 可认为是一个贪心算法。

二进制与十六进制之间的转换是非常容易的，因为每个十六进制数字对应着一组 4 个二进制数字，这种对应关系如表 2-3 所示(未表示开头的 0)。(把证明该表的正确性在本节末尾留做习题 11 和 12。)例 6 说明了这种转换。

表 2-3 整数 0 到 15 的十六进制、八进制和二进制表示

十进制	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
十六进制	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
八进制	0	1	2	3	4	5	6	7	10	11	12	13	14	15	16	17
二进制	0	1	10	11	100	101	110	111	1000	1001	1010	1011	1100	1101	1110	1111

例 6 求$(11\ 1110\ 1011\ 1100)_2$ 的十六进制展开和$(A8D)_{16}$的二进制展开。

解 为了把$(11\ 1110\ 1011\ 1100)_2$ 转化成十六进制记号，把数字分成 4 个一组，必要时在最左一组的开头加一些 0。这些组是 0011、1110、1011 和 1100，分别对应十六进制数字 3、E、B 和 C。于是$(11\ 1110\ 1011\ 1100)_2=(3EBC)_{16}$。

为了把$(A8D)_{16}$转化成二进制记号，把每个十六进制数字换成一组 4 个二进制数字。这些组是 1010、1000、1101。于是$(A8D)_{16}=(1010\ 1000\ 1101)_2$。 ■

2.5.3 整数运算算法

用整数的二进制展开做运算在计算机科学中分外重要。我们将介绍对两个表示为二进制展开的整数做加法和乘法的算法，还要以使用的位运算的实际次数来分析这些算法的计算复杂度。在整个讨论中假定 a 和 b 的二进制展开为

$$a=(a_{n-1}a_{n-2}\cdots a_1a_0)_2,\quad b=(b_{n-1}b_{n-2}\cdots b_1b_0)_2$$

从而 a 和 b 各有 n 个二进制位(必要时让其中一个的开头几位为 0)。

考虑两个用二进制记号表示的整数相加的问题。做加法的过程可以根据通常借助纸笔做加法的办法来设计，也就是成对二进制数字相加，有进位时再加上进位，这样来计算两个整数的和。现在来详细描述这个过程。

要把 a 和 b 相加，首先把最右边的数字相加。这样可得

$$a_0+b_0=c_0\cdot 2+s_0$$

其中 s_0 是 $a+b$ 的二进制展开中最右边的一位数字，而 c_0 是进位，c_0 为 0 或 1。然后把下一对二进制位及进位相加，

$$a_1+b_1+c_0=c_1\cdot 2+s_1$$

其中 s_1 是 $a+b$ 的二进制展开中的下一位(从右算起)数字，c_1 是进位。继续这一过程，把两个二进制展开中对应的二进制位及进位相加，给出 $a+b$ 的二进制展开中从右算起的下一位数字。最后把 a_{n-1}，b_{n-1}和 c_{n-2}相加得 $c_{n-1}\cdot 2+s_{n-1}$。$a+b$ 的首位数字是 $s_n=c_{n-1}$。这一过程产生 a 与 b 之和的二进制展开，即 $a+b=(s_ns_{n-1}\cdots s_1s_0)_2$。

例 7 把 $a=(1110)_2$ 和 $b=(1011)_2$ 相加。

解 按照算法中规定的步骤，首先注意到

$$a_0+b_0=0+1=0\cdot 2+1$$

所以 $c_0=0$ 而 $s_0=1$。然后，因为

$$a_1+b_1+c_0=1+1+0=1\cdot 2+0$$

所以 $c_1=1$ 而 $s_1=0$。继续做下去，

$$a_2+b_2+c_1=1+0+1=1\cdot 2+0$$

于是 $c_2=1$ 而 $s_2=0$。最后由于

$$a_3+b_3+c_2=1+1+1=1\cdot 2+1$$

从而 $c_3=1$ 且 $s_3=1$。这表明 $s_4=c_3=1$。因此 $s=a+b=(1\ 1001)_2$。相加的过程如图 2-6 所示。

```
111
 1110
 1011
-----
11001
```

图 2-6 $(1110)_2$ 和 $(1011)_2$ 相加

加法算法可用伪码描述如下。■

算法 2 整数相加

```
procedure add (a, b: 正整数)
{ a 和 b 的二进制展开分别是(a_{n-1}a_{n-2}…a_1a_0)_2 和(b_{n-1}b_{n-2}…b_1b_0)_2}
c := 0
for j := 0 to n-1
begin
    d := ⌊(a_j + b_j + c)/2⌋
    s_j := a_j + b_j + c - 2d
    c := d
end
s_n := c
{ 和数的二进制展开是(s_n s_{n-1}…s_0)_2}
```

下面分析算法 2 使用的二进制位相加的次数。

例 8 两个二进制展开中有 n 个(或少于 n 个)二进制位的整数相加，算法 2 需要用多少次二进制位加法?

解 两个整数相加是相继对二进制位相加，再加上进位来完成的。把两个二进制位及进位相加需要 3 次或少于 3 次二进制位加法。因此需要的二进制位加法总数少于二进制展开中位数的三倍。从而算法 2 把两个 n 位整数相加需要的二进制位加法次数是 $O(n)$。■

下面考虑两个 n 位整数 a 和 b 的乘法，传统的算法(用纸笔做乘法)是这样的：根据分配律看出

$$ab=a(b_0 2^0+b_1 2^1+\cdots+b_{n-1}2^{n-1})=a(b_0 2^0)+a(b_1 2^1)+\cdots+a(b_{n-1}2^{n-1})$$

可以用这一等式计算 ab。首先注意在 $b_j=1$ 时 $ab_j=a$，而 $b_j=0$ 时 $ab_j=0$。每当用 2 乘一项时，结果都是把这一项的二进制展开向左移一位并在尾部加上一个 0。因而可以把 ab_j 的二进制展开向左移位 j 位，再在尾部加上 j 个 0 来计算 $(ab_j)2^j$。最后，把 n 个整数 $ab_j 2^j$，$j=0$，1，2，…，$n-1$，相加就得到 ab。

这个乘法过程可用伪码描述成算法 3。

算法 3 整数相乘

```
procedure multiply (a, b: 正整数)
    { a 和 b 的二进制展开分别是(a_{n-1}a_{n-2}…a_1a_0)_2 和(b_{n-1}b_{n-2}…b_1b_0)_2 }
for j := 0 to n-1
begin
    if b_j = 1 then c_j := a shifted j places
    else c_j := 0
end
{ c_0, c_1, …, c_{n-1} 是部分乘积 }
p := 0
for j := 0 to n-1
    p := p + c_j
{p 是 ab 的值}
```

例 9 解释了怎样使用这一算法。

例 9 求 $a=(110)_2$ 和 $b=(101)_2$ 的乘积。

解 首先注意

$$ab_0 \cdot 2^0 = (110)_2 \cdot 1 \cdot 2^0 = (110)_2$$
$$ab_1 \cdot 2^1 = (110)_2 \cdot 0 \cdot 2^1 = (0000)_2$$

及

$$ab_2 \cdot 2^2 = (110)_2 \cdot 1 \cdot 2^2 = (11000)_2$$

为求乘积，把$(110)_2$、$(0000)_2$ 和$(11000)_2$ 相加，完成这些加法(用算法 2，必要时首位加 0)即得 $ab=(11110)_2$。这一过程如图 2-7 所示。■

```
    110
    101
    ---
    110
   000
  110
  -----
  11110
```

图 2-7 $(110)_2$ 和 $(101)_2$ 相乘

下面来判断算法 3 做乘法时使用的二进制位相加次数和移位数。

例 10 用算法 3 计算 a 和 b 的乘积需用多少次二进制位加法和移位？

解 算法 3 计算 a 和 b 之乘积的办法是把部分乘积 c_0，c_1，c_2，…，c_{n-1}相加。当 $b_j=1$ 时，部分积 c_j 的计算是把 a 的二进制展开移 j 位。当 $b_j=0$ 时，因为 $c_j=0$，所以不需要移位。于是，为求出所有 n 个整数 ab_j2^j，$j=0,1,2,\cdots,n-1$，需要至多

$$0+1+2+\cdots+n-1$$

次移位。因此根据 2.2 节例 4，需要的移位次数是 $O(n^2)$。

要把 ab_j 从 $j=0$ 到 $j=n-1$ 加起来，需要做一个 n 位整数、一个$(n+1)$位整数……和一个 $2n$ 位整数的加法。从例 8 知道，这些加法都需要 $O(n)$次二进制位相加。因此完成所有 n 个数的加法需要 $O(n^2)$次位加法。■

令人吃惊的是，还有比传统的整数乘法算法更有效的算法。其中一个算法使用 $O(n^{1.585})$次位运算来完成 n 位数的乘法，将在 6.3 节介绍。

给定整数 a 和 d，$d>0$，可以用算法 4 求 $q=a$ **div** d 和 $r=a$ **mod** d。在这个算法中，当 a 为正时，就从 a 中尽可能多次减去 d，直到剩下的值小于 d。减法的次数就是商而最后减剩下的值就是余数。算法 4 也能处理 a 为负的情况。先求出当 $|a|$ 除以 d 时的商 q 和余数 r；当 $r>0$ 时，再用这些结果求出当 a 除以 d 时的商 $-(q+1)$ 和余数 $d-r$。具体过程留给读者证明(练习 55)：假设 $a>d$，该算法用 $O(q \log a)$次位运算。

```
算法 4 计算 div 和 mod
procedure division algorithm（a：整数，d：正整数）
q := 0
r := |a|
while r≥d
begin
    r := r−d
    q := q+1
end
if a<0 且 r>0 then
begin
    r := d−r
    q := −(q+1)
end
{q=a div d 是商，r=a mod d 是余数}
```

当正整数 a 除以正整数 d 时，有比算法 4 更有效的算法能确定商 $q=a$ **div** d 和余数 $r=a$ **mod** d(详情见[Kn98])。这些算法需要 $O(\log a\cdot\log d)$次位运算。如果 a 和 d 的二进制展开都不超过 n 位，那么 $\log a\cdot\log d$ 可换成 n^2。这意味着需要 $O(n^2)$次位运算来求 a 除以 d 的商和余数。

2.5.4 同余幂

在密码学中重要的是能有效地求 b^n **mod** m，其中 b，n，m 都是大整数。先计算 b^n，再求 b^n 除以 m 的余数，这是不可行的，因为 b^n 是非常大的数。可行的是一种利用指数 n 的二进制展开(比如 $n=(a_{k-1}\cdots a_1a_0)_2$)的算法。这个算法依次求 b **mod** m，b^2 **mod** m，b^4 **mod** m，…，$b^{2^{k-1}}$ **mod** m，把其中 $a_j=1$ 的那些项 b^{2^j} **mod** m 乘起来，在每次乘法后求乘积除以 m 的余数。这个算法的伪码如算法 5 所示。

```
算法 5 同余幂
procedure modular exponentiation(b：整数，n=(a_{k-1}a_{k-2}…a_1a_0)_2，m：正整数)
x := 1
power := b mod m
for i := 0 to k−1
begin
    if a_i = 1 then x := (x · power) mod m
    power := (power · power) mod m
end
{x 等于 b^n mod m}
```

例 11 说明算法 5 如何工作。

例 11 用算法 5 求 2^{644} **mod** 645。

解 算法 5 首先令 $x=1$ 和 $power=2$ **mod** $645=2$。在计算 2^{644} **mod** 645 的过程中，这个算法对 $j=1,2,\cdots,9$ 确定 2^{2^j} **mod** 645，做法是不断地平方和模 645 缩减。如果 $a_j=1$(其中 a_j 是 644 二进制展开中的第 j 位)，就用 x 当前值乘以 2^{2^j} **mod** 645 并用模 645 来缩减结果。下面是所

用的步骤：

$i=0$：因 $a_0=0$，有 $x=1$ 和 $power=2^2=4$ **mod** 645=4；
$i=1$：因 $a_1=0$，有 $x=1$ 和 $power=4^2=16$ **mod** 645=16；
$i=2$：因 $a_2=1$，有 $x=1\cdot 16$ **mod** 645=16 和 $power=16^2=256$ **mod** 645=256；
$i=3$：因 $a_3=0$，有 $x=16$ 和 $power=256^2=65\,536$ **mod** 645=391；
$i=4$：因 $a_4=0$，有 $x=16$ 和 $power=391^2=152\,881$ **mod** 645=16；
$i=5$：因 $a_5=0$，有 $x=16$ 和 $power=16^2=256$ **mod** 645=256；
$i=6$：因 $a_6=0$，有 $x=16$ 和 $power=256^2=65\,536$ **mod** 645=391；
$i=7$：因 $a_7=1$，有 $x=(16\cdot 391)$ **mod** 645=451 和 $power=391^2=152\,881$ **mod** 645=16；
$i=8$：因 $a_8=0$，有 $x=451$ 和 $power=16^2=256$ **mod** 645=256；
$i=9$：因 $a_9=1$，有 $x=(451\cdot 256)$ **mod** 645=1。

这说明遵循算法 5 的步骤就得出结果 2^{644} **mod** 645=1。 ■

2.5.5 欧几里得算法

2.4 节介绍的用整数的素因数分解求两个整数之最大公约数的算法效率不高。原因是求素因数分解消耗太多时间。将给出一个称为欧几里得算法的效率更高的方法求最大公约数。这个算法古代就有了。这是用古希腊数学家欧几里得[⊖]的名字命名的，因为在他的书《几何原本》中记载了这一算法。

在介绍欧几里得算法之前，先说明它怎样求 gcd(91，287)。首先用两个数中的大者 287 除以两个数中的小者 91，得到

$$287 = 91\cdot 3 + 14$$

91 和 287 的任何公约数必定也是 $287-91\cdot 3=14$ 的因数。而且 91 和 14 的任何公约数也必定是 $287=91\cdot 3+14$ 的因数。因此，287 和 91 的最大公约数和 91 与 14 的最大公约数相同。这表明，求 gcd(91，287)的问题已被化简为求 gcd(91，14)的问题。

下一步用 91 除以 14 得

$$91 = 14\cdot 6 + 7$$

由于 91 和 14 的任何公约数也能整除 $91-14\cdot 6=7$，并且 14 和 7 的任何公约数整除 91，于是 gcd(91，14)=gcd(14，7)。

继续让 14 除以 7 得

$$14 = 7\cdot 2$$

由于 7 整除 14，所以 gcd(14，7)=7。另外由于 gcd(287，91)=gcd(91，14)=gcd(14，7)=7，所以最初的问题得解。

现在介绍欧几里得算法在一般情况下如何工作。将用辗转相除把求两个正整数最大公约数的问题化简为求两个较小整数的最大公约数，直到两个整数中的一个为 0。

欧几里得算法的基础是关于最大公约数的下述结论和整除算法。

⊖ **欧几里得(Euclid，约公元前 325—265 年)**是人类最成功的数学著作《几何原本》(Elments)的作者，这本书自古至今已有 1000 多个不同的版本。人们对欧几里得的生平所知甚少，只知道他曾在亚历山大著名的学院任教。显然欧几里得不强调应用。当一个学生问他学了几何有什么用时，他解释说知识本身就值得学习，并让仆人给了这个学生一枚硬币，"因为他一定要从学习中获利。"

引理 1　令 $a=bq+r$，其中 a，b，q 和 r 均为整数，则 $\gcd(a,b)=\gcd(b,r)$。

证　如果能证明 a 与 b 的公约数和 b 与 r 的公约数相同，也就证明了 $\gcd(a,b)=\gcd(b,r)$，因为这两对整数必定有相同的最大公约数。

现在假定 d 整除 a 和 b，于是 d 也整除 $a-bq=r$(根据 2.4 节定理 1)。因此 a 和 b 的任何公约数也是 b 和 r 的公约数。

类似地，假定 d 整除 b 和 r，于是 d 也整除 $bq+r=a$。因此 b 和 r 的任何公约数也是 a 和 b 的公约数。

于是 $\gcd(a,b)=\gcd(b,r)$。 □

假定 a 和 b 为正整数，$a\geqslant b$。令 $r_0=a$ 和 $r_1=b$。若辗转应用整除算法，得

$$
\begin{aligned}
r_0 &= r_1 q_1 + r_2 & 0 \leqslant r_2 < r_1 \\
r_1 &= r_2 q_2 + r_3 & 0 \leqslant r_3 < r_2 \\
&\vdots \\
r_{n-2} &= r_{n-1} q_{n-1} + r_n & 0 \leqslant r_n < r_{n-1} \\
r_{n-1} &= r_n q_n
\end{aligned}
$$

最终在辗转相除序列中会出现余数为 0，因为在余数序列 $a=r_0>r_1>r_2>\cdots\geqslant 0$ 中至多包含 a 项。进而从引理 1 知

$$
\begin{aligned}
\gcd(a,b) &= \gcd(r_0,r_1) = \gcd(r_1,r_2) = \cdots = \gcd(r_{n-2},r_{n-1}) \\
&= \gcd(r_{n-1},r_n) = \gcd(r_n,0) = r_n
\end{aligned}
$$

因此，最大公约数是除法序列中最后一个非零余数。

例 12　用欧几里得算法求 414 和 662 的最大公约数。

解　辗转使用整除算法给出：

$$
\begin{aligned}
662 &= 414 \cdot 1 + 248 \\
414 &= 248 \cdot 1 + 166 \\
248 &= 166 \cdot 1 + 82 \\
166 &= 82 \cdot 2 + 2 \\
82 &= 2 \cdot 41
\end{aligned}
$$

因此，$\gcd(414, 662)=2$，因为 2 是最后一个非零余数。 ■

算法 6 是用伪码写成的欧几里得算法。

算法 6　欧几里得算法

```
procedure gcd(a, b: 正整数)
x := a
y := b
while y≠0
begin
    r := x mod y
    x := y
    y := r
end {gcd(a, b)是 x}
```

算法 6 中 x 和 y 的初值分别是 a 和 b。在过程的每一步都是 x 取 y 的值，而 y 取 x **mod** y 的值，x **mod** y 是 x 除以 y 的余数。只要 $y\neq 0$，这个过程就重复下去。当 $y=0$ 时算法终止，此时 x

的值，也就是这一过程中最后一个非零余数，即为 a 和 b 的最大公约数。

将在3.4节研究欧几里得算法的时间复杂度，并证明求 a 和 b 的最大公约数需要做的除法次数在 $a\geqslant b$ 时为 $O(\log b)$。

练习

1. 把下列整数从十进制表示转换为二进制表示。
   a)231 b)4532 c)97644
2. 把下列整数从十进制表示转换为二进制表示。
   a)321 b)1023 c)100632
3. 把下列整数从二进制表示转换为十进制表示。
   a)1 1111 b)10 0000 0001 c)1 0101 0101 d)110 1001 0001 0000
4. 把下列整数从二进制表示转换为十进制表示。
   a)1 1011 b)10 1011 0101 c)11 1011 1110 d)111 1100 0001 1111
5. 把下列整数从十六进制表示转换为二进制表示。
   a)80E b)135AB c)ABBA d)DEFACED
6. 把 $(\mathrm{BADFACED})_{16}$ 从十六进制表示转换为二进制表示。
7. 把 $(\mathrm{ABCDEF})_{16}$ 从十六进制表示转换为二进制表示。
8. 把下列整数从二进制表示转换为十六进制表示。
   a)1111 0111 b)1010 1010 1010 c)111 0111 0111 0111
9. 把 $(1011\ 0111\ 1011)_2$ 从二进制表示转换为十六进制表示。
10. 把 $(1\ 1000\ 0110\ 0011)_2$ 从二进制表示转换为十六进制表示。
11. 证明可以从正整数的二进制展开求其十六进制展开，方法是每四个二进制位分为一组，必要时在开头加一些位，把每组四个二进制位转换成一个十六进制数字。
12. 证明可以从正整数的十六进制展开求其二进制展开，方法是把每个十六进制数字转换成四个二进制位。
13. 设计一个简单的方法把十六进制表示转换为二进制表示。
14. 设计一简单的方法把二进制表示转换为十六进制表示。
15. 把 $(7345321)_8$ 转换为二进制表示，把 $(10\ 1011\ 1011)_2$ 转换为八进制表示。
16. 给出一个从整数的十六进制表示转换为八进制表示的过程，用二进制表示作为中间步骤。
17. 给出一个从整数的八进制表示转换为十六进制表示的过程，用二进制表示作为中间步骤。
18. 把 $(12345670)_8$ 转换为十六进制表示，把 $(\mathrm{ABB093BABBA})_{16}$ 转换为八进制表示。
19. 用算法5求 $3^{2003}\ \mathbf{mod}\ 99$。
20. 用算法5求 $123^{1001}\ \mathbf{mod}\ 101$。
21. 用欧几里得算法求
   a)gcd(12, 18) b)gcd(111, 201) c)gcd(1001, 1331) d)gcd(12345, 54321)
   e)gcd(1000, 5040) f)gcd(9888, 6060)
22. 用欧几里得算法求
   a)gcd(1, 5) b)gcd(100, 101) c)gcd(123, 277) d)gcd(1529, 14039)
   e)gcd(1529, 14038) f)gcd(11111, 111111)
23. 用欧几里得算法求gcd(21, 34)要做多少次除法？
24. 用欧几里得算法求gcd(34, 55)要做多少次除法？
25. 证明每个正整数都可以唯一地表示为2的不同次幂的和。（提示：考虑整数的二进制展开。）
26. 可以证明每个整数都能表示为

$$e_k3^k+e_{k-1}3^{k-1}+\cdots+e_13+e_0$$

的形式，其中 $e_j=-1$，0 或 1，$j=0$，1，2，…，k。这一类展开称为平衡三进制展开。求下列整数的平衡三进制展开。

a)5 b)13 c)37 d)79

27. 证明正整数被 3 整除的充分必要条件是它的十进数字之和能被 3 整除。
28. 求证正整数能被 11 整除的充分必要条件是：它的偶数位十进数字之和与奇数位十进数字之和的差能被 11 整数。
29. 求证正整数能被 3 整除的充分必要条件是：它的偶数位二进数字之和与奇数位二进数字之和的差能被 3 整除。

整数的 1 补表示可以简化计算机算术。为了表示绝对值小于 2^n-1 的正负整数，一共需要 n 位。最左边一位表示符号。这一位上的 0 表示正整数，1 表示负整数。对正整数来说，其他的位就是该整数的二进制展开。对负整数来说，其他的位就是其绝对值二进制展开中各位数字的补，其中 1 以 0 为补，0 以 1 为补。

30. 用长度为六的位串，求下列整数的 1 补表示。

a)22 b)31 c)-7 d)-19

31. 下列长度为五的 1 补表示是哪些整数？

a)11001 b)01101 c)10001 d)11111

32. 如果 m 是个小于 2^{n-1} 的正整数，若用长度为 n 的位串，怎样从 m 的 1 补表示求 $-m$ 的 1 补表示？
33. 怎样从两个整数的 1 补表示计算其和的 1 补表示？
34. 怎样从两个整数的 1 补表示计算其差的 1 补表示？
35. 求证以 $(a_{n-1}a_{n-2}\cdots a_1a_0)$ 为 1 补表示的整数 m 可以用公式 $m=-a_{n-1}(2^{n-1}-1)+a_{n-2}2^{n-2}+\cdots+a_{1\cdot 2}+a_0$ 计算。

整数的 2 补表示也可以简化计算机算术，而且比 1 补表示更通用。对给定的正整数 n，要表示满足 $-2^{n-1}\leqslant x\leqslant 2^{n-1}-1$ 的整数 x，共需要 n 位。最左边一位表示符号。这一位的 0 表示正整数，1 表示负整数。这和 1 补表示一样。对正整数来说，其余的位与该整数的二进制展开一样。对负整数而言，其余位是 $2^{n-1}-|x|$ 的二进制展开。整数的 2 补表示常用于计算机运算，因为不论整数是正是负，都很容易用这种表示做加法和减法。

36. 用长度为六的位串的 2 补表示解答练习 30。
37. 假定每个表示是长度为五的 2 补表示，回答练习 31。
38. 用 2 补表示回答练习 32。
39. 用 2 补表示回答练习 33。
40. 用 2 补表示回答练习 34。
41. 证明以 $(a_{n-1}a_{n-2}\cdots a_1a_0)$ 为 2 补表示的整数 m 可以用等式 $m=-a_{n-1}\cdot 2^{n-1}+a_{n-2}\cdot 2^{n-2}+\cdots+a_1\cdot 2+a_0$ 来计算。
42. 给出一个简单的算法，从整数的 1 补表示求其 2 补表示。
43. 有时用四位二进制展开表示各个十进制数字来为整数编码。这就是整数的二进制编码十进制形式。例如，用这种方式为 791 编码得 0111 1001 0001。用这种方式为整数编码，一个 n 位的十进制数需要多少位？

所谓康托尔展开就是形如

$$a_nn!+a_{n-1}(n-1)!+\cdots+a_22!+a_11!$$

的和，其中 a_i 为整数，且 $0\leqslant a_i\leqslant i$，$i=1$，2，…，$n$。

44. 求下列各数的康托展开。

a)2 b)7 c)19 d)87 e)1000 f)1 000 000

*45. 给出一个求整数的康托表示的算法。

*46. 给出一个从两个整数的康托表示求它们的和的算法。

47. 用课文中给出的加法算法，一步一步把 $(10111)_2$ 和 $(11010)_2$ 相加。
48. 用课文中给出的乘法算法，一步一步把 $(1110)_2$ 和 $(1010)_2$ 相乘。
49. 给出一个求两个二进制展开之差的算法。
50. 估计计算两个二进制展开之差需要的位运算的次数。
51. 设计一个算法，从整数 a 和 b 已知的二进制展开判断是 $a>b$，$a=b$ 还是 $a<b$。
52. 当整数 a 和 b 的二进制展开有 n 个数位时，按练习 51 设计的比较算法要做多少次位运算?
53. 就需要的除法次数而言，估计求整数 n 的 b 进制展开的算法 1 的复杂度。
*54. 证明算法 5 用 $O((\log m)^2\log n)$ 次位运算求 $b^n \textbf{ mod } m$。
55. 证明算法 4 用 $O(q\log|a|)$ 次位运算，假设 $a>d$。

2.6 数论应用

2.6.1 引言

数论有许多应用，尤其是用于计算机科学。在 2.4 节已经介绍了几个这种应用，包括散列函数、伪随机数生成和移位密码。本节继续介绍数论，给出若干关键结果和两个重要的应用：做大整数算术的方法和最近发明的称为*公钥系统*的密码系统。在这种密码系统中，不需要为加密密钥保密，因为知道加密密钥无助于在可行的时间内解密信息。用秘密拥有的解密密钥来解密信息。

在学习这些应用之前，我们将介绍在数论及其应用中占中心地位的一些关键性结论。例如，将说明如何用中国剩余定理(即孙子定理——译者注)求解模为两两互素的整数的线性同余系统，然后说明如何以这个结果为基础做大整数算术。将介绍费尔马小定理和伪素数的概念，并说明如何用这些概念建立一个公钥系统。

2.6.2 若干有用的结果

这一节从头到尾都会用到的一个重要结果是：两个整数 a 和 b 的最大公约数可以表示为

$$sa+tb$$

的形式，其中 s 和 t 为整数。换句话说，$\gcd(a,b)$ 可以表示为 a 和 b 的整系数线性组合。例如，$\gcd(6,14)=2$，而 $2=(-2)\cdot 6+1\cdot 14$。定理 1 说的就是这一事实。

定理 1 若 a 和 b 为正整数，则存在整数 s 和 t，使 $\gcd(a,b)=sa+tb$。

将不对定理 1 作形式证明，而是举例介绍一种方法，可以求两个整数的线性组合，使之等于其最大公约数。(本节假定线性组合均以整数为系数。)这一方法是将欧几里得算法求最大公约数所做的除法依次颠倒过来。(在练习 48 前面的说明中还描述一个称为*广义欧几里得算法*的算法，可用来把 $\gcd(a,b)$ 表示成 a 和 b 的线性组合。)

例 1 把 $\gcd(252,198)=18$ 表示为 252 和 198 的线性组合。

解 要证明 $\gcd(252,198)=18$，欧几里得算法做下列除法：

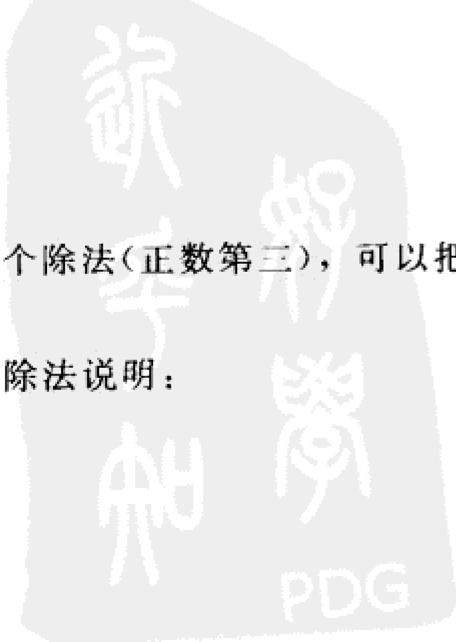
$$252=1\cdot 198+54$$
$$198=3\cdot 54+36$$
$$54=1\cdot 36+18$$
$$36=2\cdot 18$$

用倒数第二个除法(正数第三)，可以把 $\gcd(254,198)=18$ 表示为 54 和 36 的线性组合，即

$$18=54-1\cdot 36$$

第二个除法说明：

$$36=198-3\cdot 54$$



将 36 的这一表达式代入前一等式，可以把 18 表示为 54 和 198 的线性组合，也就是

$$18 = 54 - 1 \cdot 36 = 54 - 1 \cdot (198 - 3 \cdot 54) = 4 \cdot 54 - 1 \cdot 198$$

第一个除法说明：

$$54 = 252 - 1 \cdot 198$$

把 54 的这一表达式代入前面的等式，可以把 18 表示为 252 和 198 的线性组合。结论是：

$$18 = 4 \cdot (252 - 1 \cdot 198) - 1 \cdot 198 = 4 \cdot 252 - 5 \cdot 198$$

问题得解。 ■

用定理 1 推出几个有用的结果。目标之一是证明算术基本定理的一部分：每个正整数只有最多一个素因子分解。要证明：如果一个正整数有一个素数分解，而且素数是依不减的顺序写的，则这一分解是唯一的。

首先需要推导一些关于整除的结果。

引理 1 如果 a，b 和 c 为正整数，使得 $\gcd(a, b)=1$ 且 $a \mid bc$，那么 $a \mid c$。

证 由于 $\gcd(a, b)=1$，根据定理 1 知有整数 s 和 t，使

$$sa + tb = 1$$

用 c 乘等式两边，得

$$sac + tbc = c$$

由 2.4 节定理 1，可以用最后这一等式证明 $a \mid c$。根据该定理的第二部分，$a \mid tbc$。由于 $a \mid sac$ 和 $a \mid tbc$，由同一定理的第一部分可知 a 整除 $sac+tbc$，从而 $a \mid c$。这就完成了证明。 □

在证明素因子分解唯一性时，将使用引理 1 的下述推广。(引理 2 的证明在 3.3 节留做练习，因为用 3.3 节介绍的数学归纳法很容易证明。)

引理 2 如果 p 是素数，且 $p \mid a_1a_2\cdots a_n$，其中 a_i 为整数，则对于某个 i，$p \mid a_i$。

现在可以证明整数分解为素数的唯一性了，也就是证明每个整数至多有一种方式写成素数的乘积，其中素数按不减小的顺序出现。这是算术基本定理的一部分。将在 3.3 节证明另一部分，即每个整数都有素因子分解。

证 (正整数的素因子分解的唯一性)：假定正整数 n 能用两种方式写成素数的乘积，比如说 $n=p_1p_2\cdots p_s$ 和 $n=q_1q_2\cdots q_t$，其中 p_i，q_j 都是素数，而且 $p_1 \leqslant p_2 \leqslant \cdots \leqslant p_s$ 和 $q_1 \leqslant a_2 \leqslant \cdots \leqslant q_t$。

现在从两个分解式中去掉共有的素数，得

$$p_{i_1} p_{i_2} \cdots p_{i_u} = q_{j_1} q_{j_2} \cdots q_{j_v}$$

其中任何素数都不同时出现在等式两边，而 u 和 v 为正整数。由引理 2 知，有某个 k 使 p_{i_1} 整除 q_{j_k}。由于任何素数都不能整除另一个素数，所以不可能有这样的 q_{j_k}。这说明 n 至多有一种分解为素数的方式。 □

用引理 1 可以证明同余式两边除以同一个整数的一个结果。已证明(2.4 节定理 10)可以在同余式两边乘以同一个整数。但用同一个整数去除同余式两边并不一定得到有效的同余。下面的例子说明了这一点。

例 2 同余式 $14 \equiv 8 \pmod 6$ 成立，但不能两边同除以 2，因为 $14/2=7$，而 $8/2=4$，但 $7 \not\equiv 4 \pmod 6$。 ■

不过用引理 1 可以证明同余式两边可以同除以与模数互素的整数。这就是定理 2。

定理 2 令 m 为正整数，a，b 和 c 为整数。如果 $ac \equiv bc \pmod m$ 且 $\gcd(c, m)=1$，那么 $a \equiv b \pmod m$。

证 由于 $ac \equiv bc \pmod m$，则 $m \mid ac-bc=c(a-b)$。由引理 1 及 $\gcd(c, m)=1$，知 $m \mid a-b$。于是 $a \equiv b \pmod m$。 □

2.6.3　线性同余

形如 $ax\equiv b(\text{mod } m)$的同余式称为线性同余，其中 m 为正整数，a 和 b 为整数，x 为变量。在数论及其应用中到处可见这种同余。

怎样解线性同余式 $ax\equiv b(\text{mod } m)$呢？也就是说，怎样求出所有满足这一同余式的值呢？将要介绍的一个方法要利用使 $a\bar{a}\equiv 1(\text{mod } m)$的整数$\bar{a}$，如果这样的$\bar{a}$存在的话。这样的$\bar{a}$称为 a 的模 m 逆。在 a 和 m 互素的条件下，定理 3 保证 a 的模 m 逆的存在。

定理 3　如果 a 和 m 为互素的整数，$m>1$，则存在 a 的模 m 的逆。而且这个逆模 m 是唯一的。(即有小于 m 的唯一正整数$\bar{a}$，它是 a 模 m 的逆，且 a 的任何别的模 m 逆均和$\bar{a}$模 m 同余。)

证　由定理 1 及 $\gcd(a, m)=1$ 知，有整数 s 和 t 使

$$sa+tm=1$$

于是

$$sa+tm\equiv 1(\text{mod } m)$$

由于 $tm\equiv 0(\text{mod } m)$，所以

$$sa\equiv 1(\text{mod } m)$$

结论是 s 为 a 的模 m 逆。把证明这个逆是模 m 唯一的留作本节末尾的练习 9。　□

定理 3 的证明同时给出了一个在 a 和 m 互素的条件下求 a 的模 m 逆的方法：求 a 和 m 的线性组合使之等于 1(按欧几里得算法步骤倒过来做即可)；这一线性组合中 a 的系数就是 a 模 m 的一个逆。例 3 说明了这一过程。

例 3　求 3 的模 7 逆。

解　由于 $\gcd(3, 7)=1$，定理 3 说明存在 3 模 7 的逆。若用欧几里得算法求 3 和 7 的最大公约数，算法很快结束：

$$7=2\cdot 3+1$$

从这一等式看到

$$-2\cdot 3+1\cdot 7=1$$

这说明－2 是 3 模 7 的一个逆。(注意，模 7 同余于－2 的每个整数也是 3 的逆，例如 5，－9，12 等。)　■

有了 a 模 m 的逆$\bar{a}$以后，就可以很容易地解同余方程 $ax\equiv b(\text{mod } m)$；只要在线性同余式两边同乘以$\bar{a}$即可。例 4 说明了这一过程。

例 4　线性同余 $3x\equiv 4(\text{mod } 7)$的解是什么？

解　从例 3 知道－2 是 3 模 7 的逆。在同余式两边同乘以－2 得

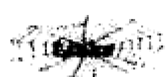
$$-2\cdot 3x\equiv -2\cdot 4(\text{mod } 7)$$

因为 $-6\equiv 1(\text{mod } 7)$且 $-8\equiv 6(\text{mod } 7)$，所以若 x 是解，必有 $x\equiv -8\equiv 6(\text{mod } 7)$。于是根据 2.4 节定理 10，有

$$3x\equiv 3\cdot 6=18\equiv 4(\text{mod } 7)$$

这说明所有这种 x 满足问题中的同余式。结论是，同余方程的解是使 $x\equiv 6(\text{mod } 7)$的整数，即 6，13，20，…及－1，－8，－15，…。　■

2.6.4　中国剩余定理



线性同余系统十分常见。例如稍后会看到，这种系统是做大整数算术的一种方式的基础。甚至可以在中国和印度的古代数学著作中发现以这种系统作为文字游戏。例 5

给出的就是这样的例子。

例 5 一世纪时，中国数学家孙子问道：

"有物不知其数，三分之余二，五分之余三，七分之余二，此物几何？"

这一问题可以翻译成：求同余方程组

$$x \equiv 2(\bmod 3)$$
$$x \equiv 3(\bmod 5)$$
$$x \equiv 2(\bmod 7)$$

的解。将在本节稍后求解这一方程组，同时也就回答了孙子的问题。 ■

从有关线性同余系统的中国古典问题而得名的中国剩余定理，可以这样叙述：只要线性同余系统的模数两两互素，则该系统有解，而且以所有模数之乘积取模，解是唯一的。

定理 4（中国剩余定理） 令 $m_1, m_2, \cdots, m_n$ 为两两互素的正整数，则同余方程组

$$x \equiv a_1(\bmod m_1)$$
$$x \equiv a_2(\bmod m_2)$$
$$\vdots$$
$$x \equiv a_n(\bmod m_n)$$

有唯一的模 $m = m_1 m_2 \cdots m_n$ 解。（即有一个解 x，使 $0 \leqslant x \leqslant m$，且所有其他的解均与此解模 m 同余。）

证 要建立这一定理，需要证明有一个解存在，而且是模 m 唯一的。下面将给出构造这样一个解的方法以证明解的存在；对这一解的模 m 唯一性的证明见本节末尾的练习 24。

要构造一个适合各方程的解，首先对 $k=1, 2, \cdots, n$，令 $M_k = m/m_k$，即 M_k 是除 m_k 以外所有模数的乘积。由于 $i \neq k$ 时，m_i 和 m_k 没有大于 1 的公因子，所以 $\gcd(m_k, M_k)=1$。从而由定理 3 知有整数 y_k，即 M_k 模 m_k 的逆，使得

$$M_k y_k \equiv 1(\bmod m_k)$$

要得到适合所有方程的解，令

$$x = a_1 M_1 y_1 + a_2 M_2 y_2 + \cdots + a_n M_n y_n$$

现在要证明 x 就是这样一个解。首先注意，由于只要 $j \neq k$，就有 $M_j \equiv 0(\bmod m_k)$，在 x 的求和表达式中除第 k 项以外的各项模 m_k 均同余于 0。由于 $M_k y_k \equiv 1(\bmod m_k)$，可看出对 $k=1, 2, \cdots, n$，均有

$$x \equiv a_k M_k y_k \equiv a_k(\bmod m_k)。$$

这就证明了 x 是这 n 个同余方程的公共解。 □

下面的例子说明怎样用定理 4 的证明中给出的构造法解同余方程组，以此来求解例 5 中孙子问题产生的方程组。

例 6 欲求解例 5 中的同余方程组，首先令 $m = 3 \cdot 5 \cdot 7 = 105$，$M_1 = m/3 = 35$，$M_2 = m/5 = 21$，$M_3 = m/7 = 15$。可看到 2 是 $M_1 = 35$ 的模 3 逆，因为 $35 \equiv 2(\bmod 3)$；1 是 $M_2 = 21$ 的模 5 逆，因为 $21 \equiv 1(\bmod 5)$；1 也是 $M_3 = 15$ 的模 7 逆，因为 $15 \equiv 1(\bmod 7)$。于是这一方程组的解是满足下列式子的那些 x：

$$\begin{aligned} x &\equiv a_1 M_1 y_1 + a_2 M_2 y_2 + a_3 M_3 y_3 = 2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 \\ &= 233 \equiv 23(\bmod 105) \end{aligned}$$

可见 23 是所有解当中最小的正整数。结论是 23 是最小的正整数，除以 3 时余 2，除以 5 时余 3，除以 7 时余 2。 ■

2.6.5 大整数计算机算术

假定 m_1，m_2，…，m_n 是大于或等于2且两两互素的整数，令 m 为它们的乘积。根据中国剩余定理可以证明(见练习22)，每个整数 $a(0\leqslant a<m)$ 均可唯一地用一个 n 元组表示，这个 n 元组由 a 除以 m_i 的余数组成，$i=1$，2，…，n。也就是说，a 可以唯一地表示为

$$(a\ \mathbf{mod}\ m_1, a\ \mathbf{mod}\ m_2, \cdots, a\ \mathbf{mod}\ m_n)$$

例7 求表示小于12的非负整数的有序对，其中第1分量是用3除的余数，第2分量是用4除的余数。

解 求出每个整数除以3和除以4的余数，得下列表示：

0=(0，0)　　4=(1，0)　　8=(2，0)

1=(1，1)　　5=(2，1)　　9=(0，1)

2=(2，2)　　6=(0，2)　　10=(1，2)

3=(0，3)　　7=(1，3)　　11=(2，3) ■

要做大整数算术，选择模数 m_1，m_2，…，m_n，其中每个 m_i 都是大于2的整数，在 $i\neq j$ 时 $\gcd(m_i, m_j)=1$，且 $m=m_1m_2\cdots m_n$ 大于要做的算术运算的结果。

一旦选定模数，大整数的算术运算就可以在表示这些整数的 n 元组的分量上做运算来完成，n 元组的分量是用大整数除以 m_i 的余数，$i=1$，2，…，n。一旦计算出表示大整数算术运算的结果的 n 元组表示，就可以求解 n 个模 m_i 同余方程($i=1$，2，…，n)找出结果的值。做大整数算术的这一方法有几个优点。首先可以用来完成通常一台计算机上不能做的大整数算术。其次，对不同的模数的计算可以并行操作，加快计算速度。

例8 假定在某台处理器上做100以内的整数算术运算比100以上的整数运算快得多，那么只要把整数表示为模两两互素的100以内的整数的余数的多元组，就可以将差不多所有整数计算限制在100以内的整数上。例如，可以用99，98，97和95为模数。(由于这些整数没有大于1的公因数，所以是两两互素的。)

根据中国剩余定理，每个小于 $99\cdot 98\cdot 97\cdot 95=89\,403\,930$ 的非负整数均可唯一地用该整数除以这四个模数的余数表示。例如把123 684表示为(33，8，9，89)，因为 $123\,684\ \mathbf{mod}\ 99=33$，$123\,684\ \mathbf{mod}\ 98=8$，$123\,684\ \mathbf{mod}\ 97=9$ 及 $123\,684\ \mathbf{mod}\ 95=89$。类似地，413 456可表示为(32，92，42，16)。

欲求123 684和413 456的和，不直接用这两个数做运算，而是使用这些四元组。我们把四元组的对应分量相加，再按相应的模数减小各个分量。这样可得

$$\begin{aligned}(33,8,9,89)+(32,92,42,16) &= (65\ \mathbf{mod}\ 99, 100\ \mathbf{mod}\ 98, 51\ \mathbf{mod}\ 97, 105\ \mathbf{mod}\ 95)\\ &= (65,2,51,10)\end{aligned}$$

要求出和，即求(65，2，51，10)表示的整数，必须解同余方程组

$$\begin{aligned}x &\equiv 65(\mathrm{mod}\ 99)\\ x &\equiv 2(\mathrm{mod}\ 98)\\ x &\equiv 51(\mathrm{mod}\ 97)\\ x &\equiv 10(\mathrm{mod}\ 95)\end{aligned}$$

可以证明(见练习39)，537 140是方程组唯一小于89 403 930的非负解，因此537 140是所求的和。注意只是在求(65，2，51，10)表示的整数时，才做大于100的整数的算术运算。■

大整数算术运算模数的最好选择是一组形为 2^k-1 的整数，其中 k 为正整数，这是因为模这

种整数的二进制算术很容易完成，且容易找到两两互素的一组这种整数。（第二个理由根据的是练习 41 证明的 $\gcd(2^a-1, 2^b-1)=2^{\gcd(a,b)}-1$ 这一事实。）例如，假定在计算机上很容易完成 2^{35} 以内的整数算术，但更大整数的运算则要求有专门的运算过程。可以使用 2^{35} 以内两两互素的一组模数做整数算术，而这些整数可以像这些模数的乘积那样大。例如，练习 42 证明了整数 $2^{35}-1, 2^{34}-1$，$2^{33}-1$，$2^{31}-1$，$2^{29}-1$ 和 $2^{23}-1$ 是两两互素的。由于这六个模数的乘积超过 2^{184}，所以可以借助对这六个不超过 2^{35} 的模数取算术模的运算完成像 2^{184} 这么大的整数的算术运算（只要运算结果也不超过这个数）。

2.6.6 伪素数

在 2.4 节证明了整数 n 为素数的条件是它不能被任何素数 p 整除，$p\leqslant\sqrt{n}$。遗憾的是，用这一标准来证明给定的整数为素数效率不高。它要求找出所有不超过 $\sqrt{n}$ 的素数，还要用这些素数试除看是否能整除 n。

有没有效率较高的方法能判断整数是否为素数呢？古代中国数学家相信，n 为素数的充分必要条件是

$$2^{n-1}\equiv 1 \pmod n$$

如果这一结论成立，就可以提供一个有效的素数检验方法。为什么他们相信这一同余式能用来判断整数是否为素数呢？首先，他们观察到只要 n 为素数，这一同余式一定成立。例如 5 是素数，而且

$$2^{5-1}=2^4=16\equiv 1 \pmod 5$$

其次，他们从未找到能使这个同余式成立的合数。这些古代中国人并非全对。他们认为只要 n 是素数这个同余式必成立，这是对的；但他们认定只要同余成立 n 就是素数，则不正确。

伟大的法国数学家费马[⊖]证明了当 n 为素数时这个同余式成立，他证明的是下述更广泛的结果。

定理 5（费马小定理） 如果 p 为素数，a 是不能被 p 整除的整数，则

$$a^{p-1}\equiv 1 \pmod p$$

因而对每个整数 a，都有

$$a^p\equiv a \pmod p$$

在本节末尾的练习 17 中给出了定理 5 的证明要点。

不幸的是，有使 $2^{n-1}\equiv 1 \pmod n$ 的合数 n 存在。这种整数称为基数为 2 的*伪素数*。

例 9 整数 341 是伪素数，因为它是合数（$341=11\cdot 31$），而且在练习 27 中证明了

$$2^{340}\equiv 1 \pmod{341}$$ ■

研究伪素数时可以使用大于 2 的整数为基数。

定义 1 令 b 是正整数。如果 n 是正合数且 $b^{n-1}\equiv 1 \pmod n$，则 n 称为*基数为 b 的伪素数*。

给定正整数 n，判断是否 $2^{n-1}\equiv 1 \pmod n$，这是一个有用的检验，可以提供关于 n 是否为素

⊖ **费马（Pierre de Fermat，1601—1665）**是 17 世纪最重要的数学家之一，从职业上来说他是位律师。他是历史上最著名的业余数学家。费马的数学发现发表得很少。我们从他与其他数学家的通信中了解他的工作。费马是解析几何的发明者之一，并且发展了微积分的某些基础思想。费马和帕斯卡为概率论建立了数学基础。费马提出了现在最有名的未解决的数学问题。他断定当 n 为大于 2 的整数时，方程 $x^n+y^n=z^n$ 没有非平凡的正整数解。300 多年来人们都没有找到证明（或反例）。在他那本古希腊数学家丢番图（Diophantus）的著作中，费马写明他有一个证明，但页边空白写不下。由于 1994 年 A·怀尔斯（Andrew Wiles）给出的第一个证明依赖复杂的现代数学，多数人认为费马自以为有了一个证明，但那证明是不正确的。不过也许是因为自己不能给出证明，他以此诱惑别人去寻找证明。

数的一些证据。具体说，如果 n 满足这个同余式，则 n 要么是素数，要么是以 2 为基数的伪素数；如果 n 不满足这个同余式，则 n 是合数。可以用 2 以外的基数 b 进行类似的检验，并获得 n 是否为素数的更多的证据。如果 n 通过所有这些检验，则 n 要么是素数，要么是以已经选择的所有 b 为基数的伪素数。另外在不超过 x 的正整数中，其中 x 是正实数，以 b 为基数的伪素数与素数比起来相对而言要少得多，其中 b 是正整数。例如，不超过 10^{10} 有 455 052 512 个素数，但只有 14 884 个以 2 为基数的伪素数。可惜的是，不能通过选择足够多的基数来区分素数与伪素数，因为有些正整数能通过满足 $\gcd(b, n)=1$ 的基数的所有检验。这样一来就有定义 2。

定义 2 对于所有满足 $\gcd(b, n)=1$ 的正整数 b，满足同余式 $b^{n-1}\equiv 1 \pmod n$ 的正合数 n 称为卡米切尔数。(这些数是用罗伯特·卡米切尔[⊖]的名字命名的，他在 20 世纪早期研究了这些数。)

例 10 整数 561 是卡米切尔数。为了看出这一点，首先注意 561 是合数，因为 $561=3\cdot 11\cdot 17$。其次注意若 $\gcd(b, 561)=1$，则 $\gcd(b, 3)=\gcd(b, 11)=\gcd(b, 17)=1$。

用费马小定理可发现

$$b^2 \equiv 1 \pmod 3),\quad b^{10} \equiv 1 \pmod{11},\quad b^{16} \equiv 1 \pmod{17}$$

所以

$$b^{560} = (b^2)^{280} \equiv 1 \pmod 3$$
$$b^{560} = (b^{10})^{56} \equiv 1 \pmod{11}$$
$$b^{560} = (b^{16})^{35} \equiv 1 \pmod{17}$$

根据本节末尾的练习 23，对于所有满足 $\gcd(b, 561)=1$ 的正整数 b 都有 $b^{560}\equiv 1 \pmod{561}$。因此 561 是卡米切尔数。 ■

尽管有无限多个卡米切尔数，可以设计更细致的检验，用做有效的概率素数性检验的基础，在本节末尾的练习中将描述这些检验。这种检验可用来迅速证明一个已知的整数几乎肯定是素数。更准确地说，如果一个整数不是素数，则这个整数通过一系列检验的概率接近于 0。第 5 章将描述这样一个检验，并讨论这个检验所依赖的一些概率论中的概念。这些概率素数性检验能够而且已经用于在计算机上迅速寻找大素数。

2.6.7 公钥密码学

在 2.4 节介绍了基于同余的信息加密法。使用这些方法时，信息(也就是字符串)被译成数字。然后每个字符对应的数用移位或模 26 的仿射变换转换为另一个数。这些方法都是私钥密码系统的例子。知道加密密钥就能迅速找到解密密钥。例如，用加密密钥 k 使用移位密码时，表示某个字母的数字 p 发送为

$$c = (p + k) \textbf{ mod } 26$$

解密是通过移位 $-k$ 来进行的，即

$$p = (c - k) \textbf{ mod } 26$$

在使用私钥密码系统时，希望秘密通信的双方必须持有各自的密钥。由于知道密钥的任何人都可以轻易地为信息加密和解密，双方必须安全交换密钥。

⊖ 罗伯特·丹尼尔·卡米切尔(Robert Daniel Carmichael, 1879—1967)罗伯特·丹尼尔·卡米切尔出生在亚拉巴马州。他于 1898 年从 Lineville 学院获得本科学位，于 1911 年从普林斯顿获得博士学位。卡米切尔于 1911～1915 年在印第安那大学，1915～1947 年在伊利诺伊大学任教。卡米切尔在广泛的领域内都是活跃的研究者，这些领域包括数论、实分析、微分方程、数学物理以及群论等。他的博士论文是在 G·D·伯克霍夫的指导下完成的，被认为是美国人对微分方程的专题所做出的第一份显著贡献。

在20世纪70年代中期，密码学者引入了公钥密码系统的概念。使用这种密码系统时，知道怎样给某人发信息并不能帮助你解密发给此人的信息。在这样一个系统中，每个人都可以有一个众所周知的加密密钥，只有解密密钥是保密的，而且只有信息的预期接收人能解密。这是因为不花费非常大量的工作(例如20亿年计算机时间)，加密密钥并不能让人找到解密密钥。

1976年，MIT的三位研究人员——李维斯特㊀、沙米尔㊁和阿德曼㊂——介绍了称为RSA系统的公钥密码系统，RSA是三个发明者姓氏的首字母缩写。RSA密码系统的基础是以两个大素数的乘积为模，求同余幂，这可用2.5节算法5来快速实现。每个人都有一把加密密钥，密钥包括模数 $n=pq$，其中 p 和 q 各自是约为200位数字的大素数，还包括与 $(p-1)(q-1)$ 互素的指数 e。要产生一把可用的密钥，必须找到两个大素数。这可以在计算机上借助本节前面提到的概率素数性检验迅速完成。这些素数的乘积 $n=pq$，大约是400位的整数，不可能在短时间内被分解。我们将看到，这正是没有一把独立的解密密钥就不可能迅速解密的重要原因。

2.6.8 RSA加密

在RSA加密法中，信息被译成若干整数序列。为此可以先将每个字母译成整数，例如用凯撒密码翻译。这些整数再分成组，各组成为一个大整数，以代表一个字母段。加密过程是先把表示普通文字(即原信息)的整数 M 转换为表示密码文字(即加密信息)的整数 C，C 的计算公式是

$$C = M^e \textbf{ mod } n$$

(为完成加密，可以使用快速的同余幂算法，例如2.5节的算法5。)加密后的信息以一段段数字的形式发送给预期的接收者。

例11说明了RSA怎样完成加密。为方便实际操作，例11中选用小素数 p 和 q，而不是100位或更大的素数。此例中描述的密码尽管并不安全，但可以解释RSA密码使用的技术。

例11 用RSA密码系统为信息STOP加密，其中 $p=43$，$q=59$，所以 $n=43\cdot 59=2537$，而且 $e=13$。注意

$$\gcd(e,(p-1)(q-1)) = \gcd(13,42\cdot 58) = 1$$

解 把STOP的字母翻译成相应的等价数字，然后按4个数字一组分段。这样得到

$$1819 \quad 1415$$

用映射

$$C = M^{13} \textbf{ mod } 2537$$

为每一段加密。用快速同余乘法算得 1819^{13} **mod** $2537=2081$ 及 1415^{13} **mod** $2537=2182$。加密后的信息为2081 2182。 ■



㊀ **李维斯特(Ronald Rivest, 1948年生)**在1969年从耶鲁获得学士学位，在1974年从斯坦福获得计算机科学博士学位。李维斯特是MIT的计算机科学教授和RSA数据安全公司的创始人，该公司拥有RSA密码系统的专利，这个系统是他与沙米尔和阿德曼一起发明的。除了密码学以外，李维斯特的研究领域有机器学习、VLSI设计和计算机算法。他是一本流行的算法教材([CoLeRiSt01])的作者之一。

㊁ **沙米尔(Adi Shamir, 1952年生)**诞生在以色列特拉维夫。他的学士学位得自特拉维夫大学(1972年)，博士学位得自魏茨曼科学研究院(1977年)。沙米尔曾任Warwick大学的助理研究员和MIT的助理教授。他现在是魏茨曼研究院应用数学系的教授，并领导一个计算机安全研究组。除了RSA以外，沙米尔对密码学的贡献有破解背包密码系统、数据加密标准(DES)的密码学分析，以及设计许多密码协议。

㊂ **阿德曼(Leonard Adleman, 1945年生)**生在加州旧金山。他从加州大学伯克利分校获得数学学士学位(1968年)和计算机科学博士学位(1976年)。阿德曼从1976年到1980年是MIT数学教职人员之一，在那里他是RSA密码系统发明人之一。他在1980年加入南加州大学(USC)计算机科学系。他1985年在USC获得讲席教职。阿德曼研究过计算机安全、计算复杂度、免疫学和分子生物学。他发明了"计算机病毒"这个术语。阿德曼最近对DNA计算的研究工作产生了极大的兴趣。他是电影《偷窥者》(Sneakers)的技术顾问。在这部电影中，计算机安全扮演了重要角色。

2.6.9 RSA 解密

当知道解密密钥 d，即 e 模 $(p-1)(q-1)$ 的逆时，就可以很快恢复原信息。(由于 $\gcd(e, (p-1)(q-1))=1$，这个逆一定存在。)要看出这一点，注意如果 $de\equiv 1(\bmod (p-1)(q-1))$，则有整数 k，使 $de=1+k(p-1)(q-1)$。由此知

$$C^d \equiv (M^e)^d = M^{de} = M^{1+k(p-1)(q-1)} (\bmod n)$$

根据费马小定理(假定 $\gcd(M, p)=\gcd(M, q)=1$，这一关系不成立的可能性非常小)，$M^{p-1}\equiv 1(\bmod p)$ 及 $M^{q-1}\equiv 1(\bmod q)$。于是

$$C^d \equiv M\cdot(M^{p-1})^{k(q-1)} \equiv M\cdot 1 \equiv M(\bmod p)$$

及

$$C^d \equiv M\cdot(M^{q-1})^{k(p-1)} \equiv M\cdot 1 \equiv M(\bmod q)$$

由于 $\gcd(p, q)=1$，从中国剩余定理知

$$C^d \equiv M(\bmod pq)$$

例 12 说明了怎样用 RSA 密码系统为信息解密。

例 12 收到的加密信息是 0981 0461。如果这是用例 11 中的 RSA 密码加密的，解密后的原信息是什么？

解 该信息是用 RSA 密码系统 $n=43\cdot 59$ 和指数 13 加密的。练习 4 证明 $d=937$ 是 13 模 $42\cdot 58=2436$ 的逆。用 937 作为解密指数。于是要解密数字块 C，计算

$$P = C^{937} \mathbf{mod}\ 2537$$

来解密上述信息，用快速同余幂算法计算 0981^{937} **mod** $2537=0704$ 及 0461^{937} **mod** $2537=1115$。从而原信息的数字码形式是 0704 1115。翻译成英文字母，原信息为 HELP。 ■

2.6.10 用 RSA 作为公钥系统

为什么 RSA 密码系统适合作为公钥密码呢？如果知道模 n 的分解，即知道素数 p 和 q，就可以用欧几里得算法迅速找到 e 模 $(p-1)(q-1)$ 的逆 d。这使得可以解密用加密密钥发送的信息。迄今还不知道有不依赖 n 的分解的解密方法。换言之，任何解密方法都要求分解 n 为素数之积。已知最有效的分解法(到 2002 年为止)需要数十亿年才能分解 400 位的整数。于是若 p 和 q 都是 200 位的素数，用 $n=pq$ 为模加密的信息，除非已知素数 p 和 q，否则不可能在合理的时间内解密。

人们正在积极研究以求发现有效分解整数的新方法。几年以前还被认为由于太大而不可能在合理的时间内分解的整数，现在已经可以较快进行分解了。超过 100 位的整数，以及若干超过 150 位的整数，已经在团队努力下分解了。一旦新的分解方法问世，就必须使用更大的素数以确保信息安全。不幸的是，先前认为安全的信息可能被并非接收方的人保存起来，然后在可以分解 RSA 加密使用的密钥中的 $n=pq$ 时解密。

RSA 方法现在得到广泛使用。不过最常用的密码系统仍是私钥密码系统。有越来越多的人使用借助 RSA 系统的公钥密码系统，不过有些应用既使用私钥又使用公钥。例如，像 RSA 这样的公钥系统可用来为希望通信的双方分配私钥。然后这些人用私钥系统来为信息加密和解密。

练习

1. 把下列各对整数的最大公约数表示为它们的线性组合。

   a)10，11　　b)21，44　　c)36，48

d)34，55　　e)117，213　　f)0，223

g)123，2347　　h)3454，4666　　i)9999，11111

2. 把下列各对整数的最大公约数表示为它们的线性组合。

a)9，11　　b)33，44　　c)35，78

d)21，55　　e)101，203　　f)124，323

g)2002，2339　　h)3457，4669　　i)10001，13422

3. 证明 15 是 7 模 26 的逆数。

4. 证明 937 是 13 模 2436 的逆数。

5. 求 4 模 9 的逆数。

6. 求 2 模 17 的逆数。

7. 求 19 模 141 的逆数。

8. 求 144 模 233 的逆数。

*9. 若 a 和 m 是互素的正整数，证明 a 模 m 的逆是模 m 唯一的。（提示：假定同余式 $ax\equiv 1 \pmod m$ 有两个解 b 和 c，再用定理 2 证明 $b\equiv c \pmod m$。）

10. 求证：若 $\gcd(a, m)>1$，则不存在 a 模 m 的逆。

11. 解同余方程 $4x\equiv 5 \pmod 9$。

12. 解同余方程 $2x\equiv 7 \pmod{17}$。

*13. 若 m 是大于 1 的正整数，$ac\equiv bc \pmod m$，求证 $a\equiv b(\bmod\ m/\gcd(c, m))$。

14. a)求证小于 11 的正整数(除 1 和 10 以外)可以分成对，每一对中的两个整数互为模 11 的逆。

b)用(a)中的结果证明 $10!\equiv -1 \pmod{11}$。

15. 求证若 p 为素数，则 $x^2\equiv 1 \pmod p$ 仅有的解是满足 $x\equiv 1 \pmod p$ 和 $x\equiv -1 \pmod p$ 的整数 x。

*16. a)推广练习 14 中(a)部分的结果，即证明若 p 为素数，则小于 p 的整数中除 1 和 $p-1$ 以外的那些可以分成对，每一对中的两个整数互为模 p 的逆。（提示：利用练习 15 中的结果。）

b)从(a)可以断定，只要 p 是素数则 $(p-1)!\equiv 1 \pmod p$。这一结果称为威尔逊(Wilson)定理。

c)若 n 为正整数，从 $(n-1)!\not\equiv 1 \pmod n$ 可以得出什么结论？

*17. 本题给出了费马小定理的一个证明轮廓。

a)假定 a 不能被素数 p 整除。证明整数 $1\cdot a, 2\cdot a, \cdots, (p-1)a$ 中的任何两个都不是模 p 同余的。

b)从(a)部分可以断定 $1, 2, \cdots, (p-1)$ 的乘积和 $a, 2a, \cdots, (p-1)a$ 的乘积是模 p 同余的。利用这一结论证明

$$(p-1)!\equiv a^{p-1}(p-1)! \pmod p$$

c)利用威尔逊定理(练习 16 中求证的)证明只要 $p\nmid a$，$a^{p-1}\equiv 1 \pmod p$。

d)利用(c)的结论证明 $a^p\equiv a \pmod p$ 对所有整数 a 成立。

18. 求同余系统的所有解。

$x\equiv 2 \pmod 3$

$x\equiv 1 \pmod 4$

$x\equiv 3 \pmod 5$

19. 求同余系统的所有解。

$x\equiv 1 \pmod 2$

$x\equiv 2 \pmod 3$

$x\equiv 3 \pmod 5$

$x\equiv 4 \pmod{11}$

*20. 求同余系统的所有解(若有解的话)。

$x\equiv 5 \pmod 6$

$x\equiv 3 \pmod{10}$

$x\equiv 8 \pmod{15}$



*21. 求同余系统的所有解(若有解的话)。

$x\equiv 7 \pmod 9$

$x\equiv 4 \pmod{12}$

$x\equiv 16 \pmod{21}$

22. 用中国剩余定理证明，任何使 $0\leqslant a<m=m_1m_2\cdots m_n$ 的整数 a 都能唯一地表示为 n 元组(a **mod** m_1，a **mod** m_2，…，a **mod** m_n)，其中 m_1，m_2，…，m_n 是两两互素的整数。

*23. 令 m_1，m_2，…，m_n 为大于或等于 2 的整数且两两互素，求证如果 $a\equiv b \pmod{m_i}$，$i=1$，2，…，n，则 $a\equiv b \pmod m$，其中 $m=m_1m_2\cdots m_n$。

*24. 证明模两两互素的一组整数的线性同余方程的公共解模这些模数之乘积是唯一的，并以此完成中国剩余定理的证明。(提示：假定 x 和 y 是两个公共解，证明对所有 i，$m_i \mid x-y$。再用练习 23 得 $m=m_1m_2\cdots m_n \mid x-y$。)

25. 哪些整数被 2 除余 1，被 3 除也余 1?

26. 哪些整数被 5 整除而被 3 除时余 1?

27. a)用费马小定理证明 $2^{340}\equiv 1 \pmod{11}$，注意 $2^{340}=(2^{10})^{34}$。

b)利用 $2^{340}=(2^5)^{68}=32^{68}$ 这一事实，证明 $2^{340}\equiv 1 \pmod{31}$。

c)从(a)和(b)推出结论 $2^{340}\equiv 1 \pmod{341}$。

28. a)用费马小定理计算 3^{302} **mod** 5，3^{302} **mod** 7 和 3^{302} **mod** 11。

b)利用(a)中结果及中国剩余定理求 3^{302} **mod** 385。(注意 $385=5\cdot 7\cdot 11$。)

29. a)用费马小定理计算 5^{2003} **mod** 7，5^{2003} **mod** 11 及 5^{2003} **mod** 13。

b)用(a)中结果及中国剩余定理求 5^{2003} **mod** 1001。(注意 $1001=7\cdot 11\cdot 13$。)

令 n 是正整数，令 $n-1=2^st$，其中 s 是非负整数，t 是正奇数。如果 $b^t\equiv 1 \pmod n$，或者对于某个 j，$0\leqslant j\leqslant s-1$，$b^{2^jt}\equiv -1 \pmod n$，就说 n 通过以 b 为基数的*米勒检验*。可以证明(见[Ro00])，一个合数 n 通过少于$n/4$个以 b 为基数的米勒检验，其中 $1<b<n$。

30. 证明若 n 是素数，b 是正整数，$n\nmid b$，则 n 通过以 b 为基数的米勒检验。

31. 证明 2047 通过以 2 为基数的米勒检验，但 2047 是合数。能通过以 b 为基数的米勒检验的正合数 n 称为*强伪素数*。于是 2047 是以 2 为基数的强伪素数。

32. 证明 1729 是卡米切尔数。

33. 证明 2821 是卡米切尔数。

*34. 证明若 $n=p_1p_2\cdots p_k$，其中 p_1，p_2，…，p_k 是不同的素数，对 $j=1$，2，…，k 满足 $p_j-1 \mid n-1$，则 n 是卡米切尔数。

35. a)用练习 34 证明当 m 是正整数，$6m+1$、$12m+1$ 和 $18m+1$ 都是素数时，每个形如$(6m+1)(12m+1)\cdot(18m+1)$的整数都是卡米切尔数。

b)用(a)证明 172 947 529 是卡米切尔数。

36. 求下列各对整数表示的小于 28 的非负整数 a，其中每对整数都表示(a **mod** 4，a **mod** 7)。

a)(0，0) b)(1，0) c)(1，1) d)(2，1) e)(2，2)

f)(0，3) g)(2，0) h)(3，5) i)(3，6)

37. 用(a **mod** 3，a **mod** 5)表示小于 15 的每一个非负整数。

38. 解释怎样用练习 37 中求出的数对计算 4 加 7。

39. 求解例 8 中得到的同余系统。

*40. 证明如果 a 和 b 为正整数，则(2^a-1) **mod** $(2^b-1)=2^{a \textbf{ mod } b}-1$。

**41. 用练习 40 证明，若 a 和 b 为正整数，那么 $\gcd(2^a-1,\ 2^b-1)=2^{\gcd(a,b)}-1$。(提示：证明用欧几里得算法计算 $\gcd(2^a-1,\ 2^b-1)$时得到的余数是形如 2^r-1 的数，其中 r 是用欧几里得算法求 $\gcd(a,\ b)$时产生的一个余数。)

42. 用练习 41 证明整数 $2^{35}-1$，$2^{34}-1$，$2^{33}-1$，$2^{31}-1$，$2^{29}-1$ 和 $2^{23}-1$ 是两两互素的。

43. 证明若 p 为奇素数，则梅森数 2^p-1 的每个因数都是形为 $2kp+1$ 的数，其中 k 是非负整数。(提示：

利用费马小定理和练习 41。）

44. 用练习 43 判断 $M_{13}=2^{13}-1=8191$ 和 $M_{23}=2^{23}-1=8\,388\,607$ 是否为素数。

*45. 证明如果知道 n 是两个素数 p 和 q 之积，且知道 $(p-1)(q-1)$ 之值，则很容易分解 n。

46. 用 RSA 系统为信息 ATTACK 加密，其中 $n=43\cdot 59$，$e=13$。把每个字母译成整数，再像例 11 那样把每一对整数算做一组。

47. 如果用 RSA 系统加密以后的信息是 0667 1947 0671，其中 RSA 的参数是 $n=43\cdot 59$，$e=13$，原信息是什么？（提示：需使用计算工具才能在合理的时间内完成计算。）

广义欧几里得算法可用来把 $\gcd(a, b)$ 表示成整数 a 和 b 的整系数线性组合。对 $j=2, 3, \cdots, n$，令 $s_0=1$，$s_1=0$，$t_0=0$ 和 $t_1=1$，令 $s_j=s_{j-2}-q_{j-1}s_{j-1}$ 和 $t_j=t_{j-2}-q_{j-1}t_{j-1}$，其中 q_j 是欧几里得算法求 $\gcd(a, b)$ 时的商（见 2.5.5 节）。可以证明（见[Ro00]）$\gcd(a, b)=s_na+t_nb$。

48. 用广义欧几里得算法把 gcd(252，356)表示成 252 和 356 的线性组合。

49. 用广义欧几里得算法把 gcd(144，89)表示成 144 和 89 的线性组合。

50. 用广义欧几里得算法把 gcd(1001，100001)表示成 1001 和 100001 的线性组合。

51. 用伪码描述广义欧几里得算法。

如果 m 是个正整数，整数 a 称为 m 的二次剩余的条件是 $\gcd(a, m)=1$ 且同余式 $x^2\equiv a \pmod m$ 有解。换言之，m 的二次剩余是与 m 互素的整数，而且是模 m 的完全平方。例如，2 是 7 的二次剩余，因为 $\gcd(2, 7)=1$，且 $3^2\equiv 2 \pmod 7$。3 不是 7 的二次剩余，因为 $\gcd(3, 7)=1$，但 $x^2\equiv 3 \pmod 7$ 无解。

52. 哪些整数是 11 的二次剩余？

53. 求证若 p 是奇素数，a 是不能被 p 整除的整数，则同余式 $x^2\equiv a \pmod p$ 要么无解，要么恰有两个模 p 不同余的解。

54. 求证若 p 是奇素数，则在 $1, 2, \cdots, p-1$ 中恰有 $(p-1)/2$ 个 p 的二次剩余。

若 p 是奇素数，a 是不能被 p 整除的整数，则勒让德符号 $\left(\frac{a}{p}\right)$ 定义为：在 a 为 p 的二次剩余时为 1，否则为 −1。

55. 求证若 p 为奇素数，a 和 b 为整数，且 $a\equiv b \pmod p$，则

$$\left(\frac{a}{p}\right)=\left(\frac{b}{p}\right)$$

56. 求证若 p 是奇素数且 a 是不能被 p 整除的正整数，则

$$\left(\frac{a}{p}\right)\equiv a^{(p-1)/2} \pmod p$$

57. 用练习 56 的结论证明若 p 是奇素数且 a 和 b 为不能被 p 整除的整数，则

$$\left(\frac{ab}{p}\right)=\left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$$

58. 证明若 p 是奇素数，则只要 $p\equiv 1 \pmod 4$，−1 就是 p 的二次剩余，若 $p\equiv 3 \pmod 4$，1 就不是 p 的二次剩余。（提示：用练习 56 的结果。）

59. 求同余式 $x^2\equiv 29 \pmod{35}$ 的所有解。（提示：求这一同余式模 5 和模 7 的解，再利用中国剩余定理。）

60. 求同余式 $x^2\equiv 16 \pmod{105}$ 的所有解。（提示：求这一同余式模 3、模 5 和模 7 的解，再利用中国剩余定理。）

2.7 矩阵

2.7.1 引言

离散数学中用矩阵表示两个集合中元素之间的关系。在随后的章节中，矩阵将用于各种不同的模型中。例如矩阵将用于通信网络模型和交通运输系统模型。许多算法都是用矩阵模型开发的。本节学习这些算法使用的矩阵算术。

定义 1 矩阵是数的矩形数组。m 行 n 列的矩阵称为 $m\times n$ 阶矩阵。(英文矩阵 matrix 的复数是 matrices。)行数和列数相同的矩阵称为方阵。若两个矩阵有同样多的行和列，而且每个位置上的对应项都相等，则这两个矩阵相等。

例 1 矩阵 $\begin{bmatrix}1 & 1\\0 & 2\\1 & 3\end{bmatrix}$ 是 3×2 矩阵。 ■

现在介绍几个矩阵术语。黑体大写字母用来表示矩阵。

定义 2 令

$$\mathbf{A}=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{bmatrix}$$

A 的第 i 行是 $1\times n$ 矩阵 $[a_{i1}, a_{i2}, \cdots, a_{in}]$。**A** 的第 j 列是 $n\times 1$ 矩阵

$$\begin{bmatrix}a_{1j}\\ a_{2j}\\ \vdots\\ a_{nj}\end{bmatrix}$$

A 的第(i, j)项或第(i, j)元素是 a_{ij}，即第 i 行第 j 列上的数。矩阵 **A** 的简写符号是 $\mathbf{A}=[a_{ij}]$，表示 **A** 是第(i, j)元素为 a_{ij} 的矩阵。

2.7.2 矩阵算术

从矩阵加法的定义开始介绍矩阵算术的基本运算。

定义 3 令 $\mathbf{A}=[a_{ij}]$ 和 $\mathbf{B}=[b_{ij}]$ 为 $m\times n$ 矩阵。**A** 和 **B** 的和用 **A**+**B** 表示，这是以 $a_{ij}+b_{ij}$ 为其第(i, j)元素的矩阵。换言之，$\mathbf{A}+\mathbf{B}=[a_{ij}+b_{ij}]$。

大小相同的两个矩阵的和是将它们对应位置上的元素相加得到的。大小不同的矩阵不能相加，因为两个矩阵的和只对行数和列数都一样的两个矩阵才有定义。

例 2 我们有

$$\begin{bmatrix}1 & 0 & -1\\2 & 2 & -3\\3 & 4 & 0\end{bmatrix}+\begin{bmatrix}3 & 4 & -1\\1 & -3 & 0\\-1 & 1 & 2\end{bmatrix}=\begin{bmatrix}4 & 4 & -2\\3 & -1 & -3\\2 & 5 & 2\end{bmatrix}$$ ■

现在讨论矩阵乘积。只有在第一个矩阵的列数和第二个矩阵的行数相等时才能定义它们的乘积。

定义 4 令 **A** 为 $m\times k$ 矩阵，**B** 为 $k\times n$ 矩阵。**A** 和 **B** 的乘积 **AB** 是个 $m\times n$ 矩阵，其第(i, j)元素等于 **A** 的第 i 行与 **B** 的第 j 列对应元素乘积之和。换言之，若 $\mathbf{AB}=[c_{ij}]$，则

$$c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{ik}b_{kj}$$

在图 2-8 中，**A** 的第 i 行和 **B** 的第 j 列用于计算 **AB** 的元素 c_{ij}。当第一个矩阵的列数和第二个矩阵的行数不相等时，不能定义它们的乘积。

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ik} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1j} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2j} & \cdots & b_{2n} \\ \vdots & & \vdots & & \vdots & \\ b_{k1} & b_{k2} & \cdots & b_{kj} & \cdots & b_{kn} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & & c_{ij} & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mn} \end{bmatrix}$$

图 2-8 $\mathbf{A}=[a_{ij}]$和$\mathbf{B}=[b_{ij}]$之积

现在举几个矩阵乘积的例子。

例 3 令

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 4 \\ 2 & 1 & 1 \\ 3 & 1 & 0 \\ 0 & 2 & 2 \end{bmatrix} \quad 和 \quad \mathbf{B} = \begin{bmatrix} 2 & 4 \\ 1 & 1 \\ 3 & 0 \end{bmatrix}$$

求 **AB**，若 **AB** 有定义。

解 因为 **A** 是 4×3 矩阵而 **B** 是 3×2 矩阵，**A** 和 **B** 的乘积有定义，是 4×2 矩阵。要计算 **AB** 的元素，首先把 **A** 的行和 **B** 的列的对应元素相乘，然后再把这些乘积加起来。例如，**AB** 的(3，1)位置的元素是 **A** 的第三行和 **B** 的第一列对应元素的乘积之和，即 $3\cdot2+1\cdot1+0\cdot3=7$。计算出 **AB** 的所有元素后，就得

$$\mathbf{AB} = \begin{bmatrix} 14 & 4 \\ 8 & 9 \\ 7 & 13 \\ 8 & 2 \end{bmatrix}$$

■

矩阵乘法是不可换的。也就是说，若 **A** 和 **B** 为矩阵，**AB** 和 **BA** 不一定相同。事实上可能这两个乘积中只有一个有定义。例如若 **A** 是 2×3 矩阵，**B** 是 3×4 矩阵，那么 **AB** 有定义，是 2×4 矩阵；**BA** 没有定义，因为 3×4 矩阵和 2×3 矩阵无法相乘。

一般来说，假定 **A** 是 $m\times n$ 矩阵，**B** 是 $r\times s$ 矩阵，则只有在 $n=r$ 时 **AB** 有定义，$s=m$ 时 **BA** 有定义。不仅如此，即使 **AB** 和 **BA** 均有定义，也只有在 $m=n=r=s$ 时 **AB** 和 **BA** 才有相同的大小，即 **AB** 和 **BA** 必定均为同样大小的方阵。再进一步说，即使 **A** 和 **B** 均为 $n\times n$ 矩阵，**AB** 和 **BA** 也不一定相等。下面就是这样的一个例子。

例 4 令

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix} \quad 和 \quad \mathbf{B} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}$$

是否 **AB**=**BA**?

解 我们发现

$$\mathbf{AB} = \begin{bmatrix} 3 & 2 \\ 5 & 3 \end{bmatrix} \quad 和 \quad \mathbf{BA} = \begin{bmatrix} 4 & 3 \\ 3 & 2 \end{bmatrix}$$

所以 **AB**≠**BA**。 ■

2.7.3 矩阵乘法算法

矩阵乘积的定义直接给出了计算两个矩阵乘积的算法。假定 $m\times n$ 矩阵 $\mathbf{C}=[c_{ij}]$是 $m\times k$ 矩阵 $\mathbf{A}=[a_{ij}]$和 $k\times n$ 矩阵 $\mathbf{B}=[b_{ij}]$的乘积。算法 1 中是用伪码表示的按矩阵乘积之定义得到的算法。

算法 1 矩阵乘法

```
procedure matrix multiplication(A, B: 矩阵)
for i := 1 to m
    for j := 1 to n
    begin
        c_ij := 0
        for q := 1 to k
            c_ij := c_ij + a_iq b_qj
    end
{C = [c_ij]是 A 和 B 的乘积}
```

可以用算法中使用的加法和乘法的次数判断这一算法的复杂度。

例 5 用算法 1 计算两个 $n\times n$ 整数矩阵的乘积，需要做多少次整数加和整数乘?

解 在 **A** 和 **B** 的乘积中有 n^2 个元素，计算每个元素要做 n 次乘法和 $n-1$ 次加法，所以一共需要 n^3 次乘法和 $n^2(n-1)$次加法。 ■

令人吃惊的是，有比算法 1 效率高的矩阵乘法算法。例 5 说明直接根据定义计算两个 $n\times n$ 矩阵的乘积需要 $O(n^3)$次乘法和加法。用别的算法计算两个 $n\times n$ 矩阵的乘积只需 $O(n^{\sqrt{7}})$次乘法和加法。(在[CoLeRiSt01]中可找到这种算法的细节。)

矩阵链乘法 涉及矩阵乘法复杂度的还有另一个重要问题，这就是怎样用尽可能少的整数乘法来计算矩阵链 $\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_n$，其中 $\mathbf{A}_1$，$\mathbf{A}_2$，…，$\mathbf{A}_n$ 分别为 $m_1\times m_2$，$m_2\times m_3$，…，$m_n\times m_{n+1}$ 整数矩阵。(本节末尾练习 13 证明矩阵乘法是可结合的，所以矩阵相乘的次序不影响计算结果。)研究这一问题之前，先要注意用算法 1 把一个 $m_1\times m_2$ 矩阵和一个 $m_2\times m_3$ 矩阵相乘需要做 $m_1m_2m_3$ 次整数乘法(参见本节末尾练习 23)。例 6 解释了这一问题。

例 6 $\mathbf{A}_1$，$\mathbf{A}_2$ 和 $\mathbf{A}_3$ 分别是 30×20，20×40 及 40×10 的整数矩阵，以什么次序计算 $\mathbf{A}_1$，$\mathbf{A}_2$ 和 $\mathbf{A}_3$ 的乘积，使用的整数乘法次数最少?

解 有两种方法计算 $\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3$ 的次序，即 $\mathbf{A}_1(\mathbf{A}_2\mathbf{A}_3)$和$(\mathbf{A}_1\mathbf{A}_2)\mathbf{A}_3$。

若 $\mathbf{A}_2$ 和 $\mathbf{A}_3$ 首先相乘，须做 $20\cdot 40\cdot 10=8000$ 次整数乘法来计算 20×10 矩阵 $\mathbf{A}_2\mathbf{A}_3$。然后需要 $30\cdot 20\cdot 10=6000$ 次乘法来计算 $\mathbf{A}_1$ 和 $\mathbf{A}_2\mathbf{A}_3$ 的乘积。因此共使用了 $8000+6000=14\,000$ 次乘法。另一方面，若 $\mathbf{A}_1$ 和 $\mathbf{A}_2$ 首先相乘，须做 $30\times 20\times 40=24\,000$ 次乘法来计算 30×40 矩阵 $\mathbf{A}_1\mathbf{A}_2$。然后需要 $30\times 40\times 10=12\,000$ 次乘法以计算 $\mathbf{A}_1\mathbf{A}_2$ 和 $\mathbf{A}_3$ 的乘积。因此共使用了 $24\,000+12\,000=36\,000$ 次乘法。

显然，第一个计算顺序更有效。 ■

[CoLeRiSt01]中讨论了判断 n 个矩阵相乘最有效方式的算法。

2.7.4 矩阵转置和幂

现在引入一个以 0 和 1 为元素的重要矩阵。

定义 5 n 阶单位阵是 $n\times n$ 矩阵 $\mathbf{I}_n=[\delta_{ij}]$，其中 $\delta_{ij}=1$ 若 $i=j$，$\delta_{ij}=0$ 若 $i\neq j$。因此

$$\mathbf{I}_n=\begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}$$

用大小合适的单位阵乘一个矩阵不改变该矩阵。换言之，若 **A** 是 $m\times n$ 矩阵，则有

$$\mathbf{AI}_n = \mathbf{I}_m\mathbf{A} = \mathbf{A}$$

可以定义方阵的幂。若 **A** 是 $n\times n$ 矩阵，则有

$$\mathbf{A}^0 = \mathbf{I}_n,\quad \mathbf{A}^r = \underbrace{\mathbf{AAA}\cdots\mathbf{A}}_{r次}$$

许多算法中都要使用把方阵的行和列交换的运算。

定义 6 令 $\mathbf{A}=[a_{ij}]$为 $m\times n$ 矩阵。**A** 的转置是交换 **A** 的行和列得到的 $n\times m$ 矩阵，用 $\mathbf{A}^t$ 表示。换言之，若 $\mathbf{A}^t=[b_{ij}]$，则 $b_{ij}=a_{ji}$，$i=1, 2, \cdots, n$，$j=1, 2, \cdots, m$。

例 7 矩阵$\begin{bmatrix}1 & 2 & 3\\4 & 5 & 6\end{bmatrix}$的转置是矩阵$\begin{bmatrix}1 & 4\\2 & 5\\3 & 6\end{bmatrix}$。 ■

行列交换之后仍是原矩阵的那些矩阵往往很重要。

定义 7 若方阵 **A** 和其转置相等，即 $\mathbf{A}=\mathbf{A}^t$，则 **A** 称为对称阵。因此 $\mathbf{A}=[a_{ij}]$为对称阵的条件是对所有 i 和 j，$1\leqslant i\leqslant n$ 和 $1\leqslant j\leqslant n$，$a_{ij}=a_{ji}$。

注意，矩阵对称的充分必要条件一是为方阵，二是对主对角线(由第 i 行第 i 列的元素组成，i 是行和列的任一序数)对称。这一对称性如图 2-9 所示。

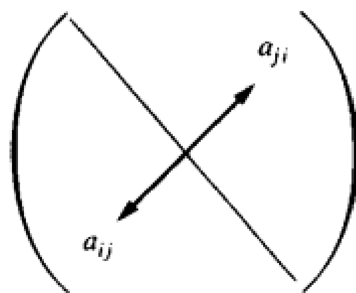


图 2-9 对称阵

例 8 矩阵$\begin{bmatrix}1 & 1 & 0\\1 & 0 & 1\\0 & 1 & 0\end{bmatrix}$是对称的。 ■

2.7.5 0-1 矩阵

元素非 0 即 1 的矩阵称为 0-1 矩阵。在第 7 章和第 8 章将会看到，0-1 矩阵常用来表示离散结构。使用这些结构的算法的基础是以 0-1 矩阵做布尔算术。布尔算术的基础是运算 ∧ 和 ∨，这些运算作用在一对二进制位上，定义如下：

$$b_1 \wedge b_2 = \begin{cases}1 & 若\ b_1 = b_2 = 1\\0 & 否则\end{cases}$$

$$b_1 \vee b_2 = \begin{cases}1 & 若\ b_1 = 1\ 或\ b_2 = 1\\0 & 否则\end{cases}$$

定义 8 令 $\mathbf{A}=[a_{ij}]$和 $\mathbf{B}=[b_{ij}]$为 $m\times n$ 阶 0-1 矩阵。**A** 和 **B** 的并，用 $\mathbf{A}\vee\mathbf{B}$ 表示，是个 0-1 矩阵，其(i, j)元素为 $a_{ij}\vee b_{ij}$。**A** 和 **B** 的交，用 $\mathbf{A}\wedge\mathbf{B}$ 表示，是个 0-1 矩阵，其(i, j)元素是$a_{ij}\wedge b_{ij}$。

例 9 求 0-1 矩阵

$$\mathbf{A}=\begin{bmatrix}1 & 0 & 1\\0 & 1 & 0\end{bmatrix},\quad \mathbf{B}=\begin{bmatrix}0 & 1 & 0\\1 & 1 & 0\end{bmatrix}$$

的并和交。

解 **A** 和 **B** 的并是

$$\mathbf{A}\vee\mathbf{B}=\begin{bmatrix}1\vee 0 & 0\vee 1 & 1\vee 0\\0\vee 1 & 1\vee 1 & 0\vee 0\end{bmatrix}=\begin{bmatrix}1 & 1 & 1\\1 & 1 & 0\end{bmatrix}$$

A 和 **B** 的交是

$$\mathbf{A}\wedge\mathbf{B}=\begin{bmatrix}1\wedge 0 & 0\wedge 1 & 1\wedge 0\\0\wedge 1 & 1\wedge 1 & 0\wedge 0\end{bmatrix}=\begin{bmatrix}0 & 0 & 0\\0 & 1 & 0\end{bmatrix}$$ ■

现在定义两个矩阵的布尔积。

定义 9 令 $\mathbf{A}=[a_{ij}]$为 $m\times k$ 阶 0-1 矩阵，$\mathbf{B}=[b_{ij}]$为 $k\times n$ 阶 0-1 矩阵。**A** 和 **B** 的布尔积用 $\mathbf{A}\odot\mathbf{B}$ 表示，是 $m\times n$ 矩阵$[c_{ij}]$，其中

$$c_{ij}=(a_{i1}\wedge b_{1j})\vee(a_{i2}\wedge b_{2j})\vee\cdots\vee(a_{ik}\wedge b_{kj})$$

注意 **A** 和 **B** 的布尔积的计算方法类似于这两个矩阵的普通乘积，但要用运算 ∨ 代替加法，用运算 ∧ 代替乘法。下面给出一个矩阵布尔乘法的例子。

例 10 求 **A** 和 **B** 的布尔积，其中

$$\mathbf{A}=\begin{bmatrix}1&0\\0&1\\1&0\end{bmatrix},\quad \mathbf{B}=\begin{bmatrix}1&1&0\\0&1&1\end{bmatrix}$$

解 **A** 和 **B** 的布尔积 $\mathbf{A}\odot\mathbf{B}$ 由下式给出：

$$\mathbf{A}\odot\mathbf{B}=\begin{bmatrix}(1\wedge1)\vee(0\wedge0)&(1\wedge1)\vee(0\wedge1)&(1\wedge0)\vee(0\wedge1)\\(0\wedge1)\vee(1\wedge0)&(0\wedge1)\vee(1\wedge1)&(0\wedge0)\vee(1\wedge1)\\(1\wedge1)\vee(0\wedge0)&(1\wedge1)\vee(0\wedge1)&(1\wedge0)\vee(0\wedge1)\end{bmatrix}$$

$$=\begin{bmatrix}1\vee0&1\vee0&0\vee0\\0\vee0&0\vee1&0\vee1\\1\vee0&1\vee0&0\vee0\end{bmatrix}$$

$$=\begin{bmatrix}1&1&0\\0&1&1\\1&1&0\end{bmatrix}$$

■

算法 2 给出了计算两个矩阵的布尔积的伪码。

```
算法 2  布尔积
procedure Boolean product(A, B: 0-1 矩阵)
for i := 1 to m
    for j := 1 to n
    begin
        c_ij := 0
        for q := 1 to k
            c_ij := c_ij ∨ (a_iq ∧ b_qj)
    end
{C=[c_ij]是 A 和 B 的布尔积}
```

还可以定义 0-1 方阵的布尔幂。这些幂将用于以后学习图中的通路，通路常用于为计算机网络中通信路径这类东西建立模型。

定义 10 令 **A** 为 0-1 方阵，r 为正整数。**A** 的 r 次布尔幂是 r 个 **A** 的布尔积。**A** 的 r 次布尔幂用 $\mathbf{A}^{[r]}$表示，因此

$$\mathbf{A}^{[r]}=\underbrace{\mathbf{A}\odot\mathbf{A}\odot\mathbf{A}\odot\cdots\odot\mathbf{A}}_{r\text{次}}$$

(由于矩阵的布尔积是可结合的，这一定义是合理的。)还定义 $\mathbf{A}^{[0]}$为 $\mathbf{I}_n$。

例 11 令 $\mathbf{A}=\begin{bmatrix}0&0&1\\1&0&0\\1&1&0\end{bmatrix}$。对所有正整数 n 求 $\mathbf{A}^{[n]}$。

解 可以求出

$$\mathbf{A}^{[2]}=\mathbf{A}\odot\mathbf{A}=\begin{bmatrix}1&1&0\\0&0&1\\1&0&1\end{bmatrix}$$

还可以求出

$$\mathbf{A}^{[3]}=\mathbf{A}^{[2]}\odot\mathbf{A}=\begin{bmatrix}1&0&1\\1&1&0\\1&1&1\end{bmatrix},\mathbf{A}^{[4]}=\mathbf{A}^{[3]}\odot\mathbf{A}=\begin{bmatrix}1&1&1\\1&0&1\\1&1&1\end{bmatrix}$$

继续计算证明

$$\mathbf{A}^{[5]}=\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$$

读者现在可以看出 $\mathbf{A}^{[n]}=\mathbf{A}^{[5]}$，对所有正整数 n，$n\geqslant 5$。 ■

很容易求出计算两个 $n\times n$ 矩阵的布尔积需要的位运算次数。

例 12 若 **A** 和 **B** 为 $n\times n$ 阶 0-1 矩阵，计算 $\mathbf{A}\odot\mathbf{B}$ 需要做多少次位运算？

解 $\mathbf{A}\odot\mathbf{B}$ 中有 n^2 个元素。用算法 2，需要 n 次 $\vee$ 和 n 次 $\wedge$ 来计算 $\mathbf{A}\odot\mathbf{B}$ 的一个元素。因此每求一个元素需要 $2n$ 次位运算。所以用算法 2 计算 $\mathbf{A}\odot\mathbf{B}$ 需要 $2n^3$ 次位运算。 ■

练习

1. 令 $\mathbf{A}=\begin{bmatrix}1&1&1&3\\2&0&4&6\\1&1&3&7\end{bmatrix}$。
   a) **A** 的大小是什么？
   b) **A** 的第 3 列是什么？
   c) **A** 的第 2 行是什么？
   d) **A** 在(3，2)位置上的元素是什么？
   e) $\mathbf{A}^t$ 是什么？
2. 求 **A**+**B**，其中
   a) $\mathbf{A}=\begin{bmatrix}1&0&4\\-1&2&2\\0&-2&-3\end{bmatrix}$，$\mathbf{B}=\begin{bmatrix}-1&3&5\\2&2&-3\\2&-3&0\end{bmatrix}$
   b) $\mathbf{A}=\begin{bmatrix}-1&0&5&6\\-4&-3&5&-2\end{bmatrix}$，$\mathbf{B}=\begin{bmatrix}-3&9&-3&4\\0&-2&-1&2\end{bmatrix}$
3. 求 **AB**，若
   a) $\mathbf{A}=\begin{bmatrix}2&1\\3&2\end{bmatrix}$，$\mathbf{B}=\begin{bmatrix}0&4\\1&3\end{bmatrix}$
   b) $\mathbf{A}=\begin{bmatrix}1&-1\\0&1\\2&3\end{bmatrix}$，$\mathbf{B}=\begin{bmatrix}3&-2&-1\\1&0&2\end{bmatrix}$
   c) $\mathbf{A}=\begin{bmatrix}4&-3\\3&-1\\0&-2\\-1&5\end{bmatrix}$，$\mathbf{B}=\begin{bmatrix}-1&3&2&-2\\0&-1&4&-3\end{bmatrix}$
4. 求乘积 **AB**，其中

a) $\mathbf{A}=\begin{bmatrix}1&0&1\\0&-1&-1\\-1&1&0\end{bmatrix}$， $\mathbf{B}=\begin{bmatrix}0&1&-1\\1&-1&0\\-1&0&1\end{bmatrix}$

b) $\mathbf{A}=\begin{bmatrix}1&-3&0\\1&2&2\\2&1&-1\end{bmatrix}$， $\mathbf{B}=\begin{bmatrix}1&-1&2&3\\-1&0&3&-1\\-3&-2&0&2\end{bmatrix}$

c) $\mathbf{A}=\begin{bmatrix}0&-1\\7&2\\-4&-3\end{bmatrix}$， $\mathbf{B}=\begin{bmatrix}4&-1&2&3&0\\-2&0&3&4&1\end{bmatrix}$

5. 求矩阵 **A**，使得 $\begin{bmatrix}2&3\\1&4\end{bmatrix}\mathbf{A}=\begin{bmatrix}3&0\\1&2\end{bmatrix}$。（提示：求 **A** 需要解线性方程组。）

6. 求矩阵 **A**，使得 $\begin{bmatrix}1&3&2\\2&1&1\\4&0&3\end{bmatrix}\mathbf{A}=\begin{bmatrix}7&1&3\\1&0&3\\-1&-3&7\end{bmatrix}$。

7. 令 **A** 为 $m\times n$ 矩阵，**O** 为元素全为 0 的 $m\times n$ 矩阵，证明 $\mathbf{A}=\mathbf{A}+\mathbf{O}=\mathbf{O}+\mathbf{A}$。

8. 证明矩阵加法是可换的，即证明若 **A** 和 **B** 均为 $m\times n$ 矩阵，则 $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$。

9. 证明矩阵加法是可结合的，即证明若 **A**，**B** 和 **C** 均为 $m\times n$ 矩阵，则 $\mathbf{A}+(\mathbf{B}+\mathbf{C})=(\mathbf{A}+\mathbf{B})+\mathbf{C}$。

10. 令 **A** 为 3×4 矩阵，**B** 为 4×5 矩阵，**C** 是 4×4 矩阵。判断下列哪些乘积有定义，求出有定义的那些乘积的大小。

    a) **AB**　　b) **BA**　　c) **AC**　　d) **CA**　　e) **BC**　　f) **CB**

11. 如果乘积 **AB** 和 **BA** 均有定义，对矩阵 **A** 和 **B** 的大小能知道些什么？

12. 本题要证明矩阵乘法对矩阵加法有分配律。

    a) 假定 **A**，**B** 均为 $m\times k$ 矩阵，**C** 为 $k\times n$ 矩阵，求证 $(\mathbf{A}+\mathbf{B})\mathbf{C}=\mathbf{AC}+\mathbf{BC}$。

    b) 假定 **C** 是 $m\times k$ 矩阵，**A** 和 **B** 为 $k\times n$ 矩阵，求证 $\mathbf{C}(\mathbf{A}+\mathbf{B})=\mathbf{CA}+\mathbf{CB}$。

13. 本题要证明矩阵乘法有结合律。假定 **A** 是 $m\times p$ 矩阵，**B** 是 $p\times k$ 矩阵。**C** 是 $k\times n$ 矩阵。求证 $\mathbf{A}(\mathbf{BC})=(\mathbf{AB})\mathbf{C}$。

14. $n\times n$ 矩阵 $\mathbf{A}=[a_{ij}]$ 称为*对角阵*，若对所有 $i\neq j$，$a_{ij}=0$。求证两个 $n\times n$ 对角阵的乘积仍是对角阵。给出一个计算这一乘积的简单规则。

15. 令 $\mathbf{A}=\begin{bmatrix}1&1\\0&1\end{bmatrix}$，求计算 $\mathbf{A}^n$ 的公式，其中 n 为正整数。

16. 证明 $(\mathbf{A}^t)^t=\mathbf{A}$。

17. 令 **A** 和 **B** 为两个 $n\times n$ 矩阵，求证

    a) $(\mathbf{A}+\mathbf{B})^t=\mathbf{A}^t+\mathbf{B}^t$　　b) $(\mathbf{AB})^t=\mathbf{B}^t\mathbf{A}^t$

若 **A** 和 **B** 是 $n\times n$ 矩阵且 $\mathbf{AB}=\mathbf{BA}=\mathbf{I}_n$，则称 **B** 为 **A** 的*逆*（这一术语是适当的，因为这样的 **B** 是唯一的），而 **A** 是*可逆的*。记号 $\mathbf{B}=\mathbf{A}^{-1}$ 表示 **B** 是 **A** 的逆。

18. 证明 $\begin{bmatrix}2&3&-1\\1&2&1\\-1&-1&3\end{bmatrix}$ 是 $\begin{bmatrix}7&-8&5\\-4&5&-3\\1&-1&1\end{bmatrix}$ 的逆。

19. 令 **A** 为 2×2 矩阵，$\mathbf{A}=\begin{bmatrix}a&b\\c&d\end{bmatrix}$。证明若 $ad-bc\neq 0$，则 $\mathbf{A}^{-1}=\begin{bmatrix}\dfrac{d}{ad-bc}&\dfrac{-b}{ad-bc}\\\dfrac{-c}{ad-bc}&\dfrac{a}{ad-bc}\end{bmatrix}$。

20. 令 $\mathbf{A}=\begin{bmatrix}-1&2\\1&3\end{bmatrix}$。

    a) 求 $\mathbf{A}^{-1}$（提示：用练习 19 的结果）。

    b) 求 $\mathbf{A}^3$。

c)求$(\mathbf{A}^{-1})^3$。

d)用(b)和(c)的答案证明$(\mathbf{A}^{-1})^3=(\mathbf{A}^3)^{-1}$。

21. 令 **A** 为可逆矩阵。证明只要 n 是正整数，就有$(\mathbf{A}^n)^{-1}=(\mathbf{A}^{-1})^n$。

22. 令 **A** 为矩阵。求证 $\mathbf{AA}^t$ 是对称的。(提示：借助练习 17b，证明这一矩阵等于其转置。)

23. 证明传统算法计算 $m_1\times m_2$ 矩阵 **A** 和 $m_2\times m_3$ 矩阵 **B** 的乘积需要 $m_1m_2m_3$ 次乘法。

24. 计算矩阵 $\mathbf{A}_1$，$\mathbf{A}_2$ 和 $\mathbf{A}_3$ 之乘积的最有效方法是什么？$\mathbf{A}_1$，$\mathbf{A}_2$，$\mathbf{A}_3$ 的大小是

a)20×50，50×10，10×40　　b)10×5，5×50，50×1

25. 把 $\mathbf{A}_1$，$\mathbf{A}_2$，$\mathbf{A}_3$ 和 $\mathbf{A}_4$ 相乘的最有效方法是什么？其中 $\mathbf{A}_1$，$\mathbf{A}_2$，$\mathbf{A}_3$ 和 $\mathbf{A}_4$ 的大小分别是10×2,2×5，5×20，20×3。

26. a)求证以 x_1，x_2，…，x_n 为变量的线性方程组

$$\begin{aligned}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n&=b_1\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n&=b_2\\&\vdots\\a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n&=b_n\end{aligned}$$

可以表示为 $\mathbf{AX}=\mathbf{B}$，其中 $\mathbf{A}=[a_{ij}]$，**X** 是 $n\times1$ 矩阵，且 x_i 就是其第 i 行，**B** 是 $n\times1$ 矩阵，b_i 是其第 i 行。

b)求证若矩阵 $\mathbf{A}=[a_{ij}]$是可逆的(在练习 18 前面定义了可逆)，则(a)中方程组的解可以用等式 $\mathbf{X}=\mathbf{A}^{-1}\mathbf{B}$ 计算。

27. 用练习 18 和 26 解方程组

$$\begin{aligned}7x_1-8x_2+5x_3&=5\\-4x_1+5x_2-3x_3&=-3\\x_1-x_2+x_3&=0\end{aligned}$$

28. 令 $\mathbf{A}=\begin{bmatrix}1&1\\0&1\end{bmatrix}$和 $\mathbf{B}=\begin{bmatrix}0&1\\1&0\end{bmatrix}$。求 a)$\mathbf{A}\vee\mathbf{B}$；b)$\mathbf{A}\wedge\mathbf{B}$；c)$\mathbf{A}\odot\mathbf{B}$。

29. 令 $\mathbf{A}=\begin{bmatrix}1&0&1\\1&1&0\\0&0&1\end{bmatrix}$和 $\mathbf{B}=\begin{bmatrix}0&1&1\\1&0&1\\1&0&1\end{bmatrix}$。求 a)$\mathbf{A}\vee\mathbf{B}$；b)$\mathbf{A}\wedge\mathbf{B}$；c)$\mathbf{A}\odot\mathbf{B}$。

30. 求 **A** 和 **B** 的布尔积，其中 $\mathbf{A}=\begin{bmatrix}1&0&0&1\\0&1&0&1\\1&1&1&1\end{bmatrix}$和 $\mathbf{B}=\begin{bmatrix}1&0\\0&1\\1&1\\1&0\end{bmatrix}$。

31. 令 $\mathbf{A}=\begin{bmatrix}1&0&0\\1&0&1\\0&1&0\end{bmatrix}$。求 a)$\mathbf{A}^{[2]}$；b)$\mathbf{A}^{[3]}$；c)$\mathbf{A}\vee\mathbf{A}^{[2]}\vee\mathbf{A}^{[3]}$。

32. 令 **A** 为 0-1 矩阵。求证 a)$\mathbf{A}\vee\mathbf{A}=\mathbf{A}$；b)$\mathbf{A}\wedge\mathbf{A}=\mathbf{A}$。

33. 本题要证明交和并运算是可换的。令 **A** 和 **B** 为 $m\times n$ 阶 0-1 矩阵。求证

a)$\mathbf{A}\vee\mathbf{B}=\mathbf{B}\vee\mathbf{A}$　　b)$\mathbf{B}\wedge\mathbf{A}=\mathbf{A}\wedge\mathbf{B}$

34. 本题要证明交和并运算是可结合的。令 **A**，**B** 和 **C** 为 $m\times n$ 阶 0-1 矩阵。求证

a)$(\mathbf{A}\vee\mathbf{B})\vee\mathbf{C}=\mathbf{A}\vee(\mathbf{B}\vee\mathbf{C})$　　b)$(\mathbf{A}\wedge\mathbf{B})\wedge\mathbf{C}=\mathbf{A}\wedge(\mathbf{B}\wedge\mathbf{C})$

35. 本题要建立交对并运算的分配律。令 **A**，**B** 和 **C** 为 $m\times n$ 阶 0-1 矩阵。求证

a)$\mathbf{A}\vee(\mathbf{B}\wedge\mathbf{C})=(\mathbf{A}\vee\mathbf{B})\wedge(\mathbf{A}\vee\mathbf{C})$　　b)$\mathbf{A}\wedge(\mathbf{B}\vee\mathbf{C})=(\mathbf{A}\wedge\mathbf{B})\vee(\mathbf{A}\wedge\mathbf{C})$

36. 令 **A** 为 $n\times n$ 阶 0-1 矩阵，令 **I** 为 $n\times n$ 单位阵。求证 $\mathbf{A}\odot\mathbf{I}=\mathbf{I}\odot\mathbf{A}=\mathbf{A}$。

37. 本题要证明 0-1 矩阵的布尔积是可结合的。假定 **A** 是 $m\times p$ 阶 0-1 矩阵，**B** 是 $p\times k$ 阶 0-1 矩阵，**C** 是 $k\times n$ 阶0-1 矩阵。求证 $\mathbf{A}\odot(\mathbf{B}\odot\mathbf{C})=(\mathbf{A}\odot\mathbf{B})\odot\mathbf{C}$。

关键术语和结果

术语

算法：一组有限多条准确的指令，用于完成一个计算或解答一个问题。
搜索算法：为元素在表中定位的问题。
线性搜索算法：逐个搜索表中元素的过程。
二分搜索算法：逐次对半搜索有序表的过程。
排序：按非升顺序重新排列表中元素。
贪心算法：每一步都做最好选择的算法。
$f(x)$是$O(g(x))$：对某个常数C和k，对所有$x>k$，$|f(x)|\leq C|g(x)|$的事实。
$f(x)$是$O(g(x))$这一关系的见证：一对C和k，使得只要$x>k$，就有$|f(x)|\leq C|g(x)|$。
$f(x)$是$\Omega(g(x))$：对某个正常数C和k，对所有$x>k$，$|f(x)|\geq C|g(x)|$的事实。
$f(x)$是$\Theta(g(x))$：$f(x)$是$O(g(x))$和$f(x)$是$\Omega(g(x))$的事实。
时间复杂度：算法解题需要的时间量。
空间复杂度：算法解题需要的存储空间量。
最坏情形时间复杂度：算法解答一定大小的问题需要的最大时间量。
平均情形时间复杂度：算法解答一定大小的问题需要的平均时间量。
$a\mid b$（a能整除b）：有整数c使$b=ac$。
素数：大于1且恰有两个正整数因子的正整数。
合数：大于1又不是素数的正整数。
梅森素数：形如2^p-1的素数，其中p为素数。
$\gcd(a,b)$（a和b的最大公约数）：能整除a和b的最大整数。
互素整数：使$\gcd(a,b)=1$的整数a和b。
两两互素的整数：一组整数，其中任何两个整数都互素。
$\mathrm{lcm}(a,b)$（a和b的最小公倍数）：能被a和b整除的最小正整数。
$a \bmod b$：a除以正整数b时的余数。
$a\equiv b \pmod m$（a模m同余b）：$a-b$能被m整除。
加密：使信息成为秘密的过程。
解密：将加密信息还原的过程。
$n=(a_k a_{k-1}\cdots a_1 a_0)_b$：$n$的$b$进制表示。
二进制表示：整数以2为基数的表示。
十六进制表示：整数以16为基数的表示。
八进制表示：整数以8为基数的表示。
a和b的整系数线性组合：形如$sa+tb$的数，其中s和t为整数。
a模m的逆：使$\bar{a}a\equiv 1 \pmod m$的整数$\bar{a}$。
线性同余：形如$ax\equiv b \pmod m$的同余式，其中x为变量。
以2为基的伪素数：使$2^{n-1}\equiv 1 \pmod n$的合数n。
以b为基的伪素数：使$b^{n-1}\equiv 1 \pmod n$的合数n。
卡米切尔数：合数n使得对所有满足$\gcd(b,n)=1$的正整数b，n是以b为基的伪素数。
私钥加密：加密密钥和解密密钥均须保密的加密法。

公钥加密：加密密钥公开，解密密钥保密的加密法。

矩阵：数的矩形数组。

矩阵加法：见 2.7 节定义 3。

矩阵乘法：见 2.7 节定义 4。

$\mathbf{I}_n$（n 阶单位阵）：对角线元素为 1、其他元素为 0 的 $n\times n$ 阶矩阵。

$\mathbf{A}^t$（$\mathbf{A}$ 的转置）：交换 $\mathbf{A}$ 的行和列得到的矩阵。

对称阵：若一个矩阵与其转置相等，这个矩阵是对称的。

0-1 矩阵：元素非 0 即 1 的矩阵。

$\mathbf{A}\vee\mathbf{B}$（$\mathbf{A}$ 和 $\mathbf{B}$ 的并）：见 2.7 节定义 8。

$\mathbf{A}\wedge\mathbf{B}$（$\mathbf{A}$ 和 $\mathbf{B}$ 的交）：见 2.7 节定义 8。

$\mathbf{A}\odot\mathbf{B}$（$\mathbf{A}$ 和 $\mathbf{B}$ 的布尔积）：见 2.7 节定义 9。

结果

线性和二分搜索算法：（见 2.1 节）。

冒泡排序：在每一遍都交换错位的相邻元素的排序方法。

插入排序：在第 j 步把第 j 个元素插入前 $j-1$ 个元素已排序的表的正确位置的排序方法。

线性搜索有 $O(n)$ 复杂度。

二分搜索有 $O(\log n)$ 复杂度。

冒泡和插入排序有 $O(n^2)$ 复杂度。

$\log n!$ 是 $O(n\log n)$。

若 $f_1(x)$ 是 $O(g_1(x))$ 且 $f_2(x)$ 是 $O(g_2(x))$，则 $(f_1+f_2)(x)$ 是 $O(\max(g_1(x), g_2(x)))$ 且 $(f_1f_2)(x)$ 是 $O((g_1(x)g_2(x))$。

若 $a_0, a_1, \cdots, a_n$ 都是实数且 $a_n\neq 0$，则 $a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$ 是 $O(x^n)$ 和 $\Theta(x^n)$。

算术基本定理：每个正整数均可唯一表示为素数之积，其中的素因子按从小到大的次序出现。

整除算法：令 a 和 d 为整数，d 为正整数，则有唯一的整数 q 和 r，$0\leqslant r<d$，使 $a=dq+r$。

如果 a，b 为正整数，则 $ab=\gcd(a, b)\cdot\operatorname{lcm}(a, b)$。

欧几里得算法：用于求最大公约数（见 2.5 节算法 6）。

令 b 是大于 1 的正整数。只要 n 是正整数，n 就能唯一表示为 $n=a_kb^k+a_{k-1}b^{k-1}+\cdots+a_1b+a_0$ 的形式。

求整数以 b 为基数的展开式的算法（见 2.5 节算法 1）。

传统的整数加法和乘法算法（见 2.5 节）。

同余幂算法（见 2.5 节算法 5）。

两个整数的最大公约数可以表示为这两个整数的整系数线性组合。

如果 m 是正整数且 $\gcd(a, m)=1$，则 a 有唯一的模 m 逆。

中国剩余定理（孙子定理）：以一组两两互素的整数为模的线性同余系统有以这些数之积为模的唯一解。

费马小定理：若 p 为素数且 $p\nmid a$，则 $a^{p-1}\equiv 1\pmod{p}$。

复习题

1. a）定义术语算法。

   b）有什么不同的方式描述算法？

c)解一个问题的算法和解同样问题的计算机程序有什么不同?

2. a)用中文描述一个求 n 个整数的表中最大整数的算法。
   b)用伪码表达这一算法。
   c)这个算法使用多少次比较?

3. a)叙述一下 $f(n)$ 是 $O(g(n))$ 这一事实的定义，其中 $f(n)$ 和 $g(n)$ 是从正整数集到实数集的函数。
   b)直接用 $f(n)$ 是 $O(g(n))$ 这一事实的定义，求证或反驳 $n^2+18n+107$ 是 $O(n^3)$。
   c)直接用 $f(n)$ 是 $O(g(n))$ 这一事实的定义，求证或反驳 n^3 是 $O(n^2+18n+107)$。

4. a)如何得出对一个函数的大 O 估计，这个函数是一些不同项之和，其中每项又是几个函数之积?
   b)给出函数 $f(n)=(n!+1)(2^n+1)+(n^{n-2}+8n^{n-3})(n^3+2^n)$ 的大 O 估计。对 $f(x)$ 是 $O(g(x))$ 的估计中的函数 g，要用次数最低的最简单的函数。

5. a)对一个求一串 n 个整数中最小整数的算法，给出最坏情形时间复杂度、平均情形时间复杂度和最好情形时间复杂度的含义(用比较次数来度量)。
   b)将当前已找到的最小整数与每个整数比较以求一串 n 个整数中最小整数的算法，若用比较次数来度量，其最坏情形、平均情形和最好情形时间复杂度是什么?

6. a)描述一下在一串以增序排列的整数中找一个整数的线性搜索和二分搜索算法。
   b)比较上述两个算法的最坏情形时间复杂度。
   c)(用比较次数来度量)是否这两个算法其中一个总是比另一个快?

7. a)描述一下冒泡排序算法。
   b)用冒泡排序算法为表 5，2，4，1，3 排序。
   c)给出对冒泡排序算法所用比较次数的大 O 估计。

8. a)描述一下插入排序算法。
   b)用插入排序算法为表 2，5，1，4，3 排序。
   c)给出对插入排序算法所用比较次数的大 O 估计。

9. a)解释一下贪心算法的概念。
   b)给出一个求出最优解的贪心算法的例子，并解释为什么求出最优解。
   c)给出一个不总是求出最优解的贪心算法的例子，并解释为什么不能求出最优解。

10. 叙述算术基本定理。

11. a)描述一个将整数分解为素数因子的过程。
    b)用这一过程分解 80 707 为素因子。

12. a)定义两个整数的最大公约数。
    b)给出至少三种求两个整数最大公约数的方法，每种方法在什么情况下最有效?
    c)求 1 234 567 和 7 654 321 的最大公约数。
    d)求 $2^3 3^5 5^7 7^9 11$ 和 $2^9 3^7 5^5 7^3 13$ 的最大公约数。

13. a)定义什么是 a 和 b 模 7 同余。
    b)整数 −11，−8，−7，−1，0，3 和 17 中，哪几对整数模 7 同余?
    c)证明若 a 和 b 模 7 同余，则 $10a+13$ 和 $-4b+20$ 也模 7 同余。

14. 描述一个将整数的十进制(以 10 为基数)表示转换为十六进制表示的过程。

15. a)如何求两个整数的(整系数)线性组合，使之等于其最大公约数?
    b)把 gcd(84，119)表达为 84 和 119 的线性组合。

16. a) $\bar{a}$ 为 a 模 m 的逆是什么意思?
    b)若 m 是正整数且 $\gcd(a, m)=1$，怎样求 a 模 m 的逆?
    c)求 7 模 19 的逆。

17. a)在 $\gcd(a, m)=1$ 时，怎样用 a 模 m 的逆求解线性同余方程 $ax\equiv b \pmod{m}$?
    b)解线性同余方程 $7x\equiv 13 \pmod{19}$。

18. a)叙述中国剩余定理。

b)求同余式 $x\equiv 1 \pmod 4$，$x\equiv 2 \pmod 5$ 和 $x\equiv 3 \pmod 7$ 的解。

19. 假定 $2^{n-1}\equiv 1 \pmod n$。n 一定是素数吗？

20. a)私钥密码系统和公钥密码系统的区别是什么？
    b)解释一下为什么移位密码是私钥系统。
    c)解释一下为什么 RSA 密码系统是公钥系统。

21. 定义两个矩阵 **A** 和 **B** 的乘积。这一乘积何时有定义？

22. a)在矩阵乘积 $\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3\mathbf{A}_4$ 有定义时，通过依次做两个矩阵的乘法，共有多少种不同方式计算这一乘积？
    b)假定 $\mathbf{A}_1$，$\mathbf{A}_2$，$\mathbf{A}_3$ 和 $\mathbf{A}_4$ 分别为 10×20，20×5，5×10 和 10×5 矩阵。怎样才能用最少次数的元素相乘来计算乘积 $\mathbf{A}_1\mathbf{A}_2\mathbf{A}_3\mathbf{A}_4$？

补充练习

1. a)描述一个算法，为一列整数中最大整数的最后一次出现定位。
   b)估计一下所用的比较次数。

2. a)描述一个算法，求一列整数中的最大和次大整数。
   b)估计一下所用的比较次数。

3. a)给出一个算法，判断一个位串中是否含两个相邻的 0。
   b)这个算法用多少次比较？

4. a)假定一列整数按从大到小的次序排列，而且整数可以重复出现。设计一个算法，求整数 x 在这个整数列中的所有出现位置。
   b)估计一下所用的比较次数。

5. a)修改 2.1 节算法 1，求一列 n 个元素中的最大和最小元素，方法是检查相邻元素时，更新临时最大和临时最小元素。
   b)用伪码描述(a)中算法。
   c)这个算法要比较多少次序列中的元素？(不算用来判断是否到达序列结尾的那些比较。)

6. a)详细描述一个算法的步骤，这个算法求一列 n 个元素中的最大和最小元素，方法是检查相邻元素对，记录下临时最大和临时最小元素。若 n 是奇数，则临时最大和临时最小元素在开始时都等于第一个元素。若 n 是偶数，则应通过比较前两个元素来找出临时最大和临时最小元素。通过与所检查的两个元素的最大和最小元素相比较，来更新临时最大和临时最小元素。
   b)用伪码描述(a)中算法。
   c)这个算法要比较多少次序列中的元素？(不算用来判断是否到达序列结尾的那些比较。)

*7. 求证就比较次数而言，求 n 个元素中最大和最小元素的算法的最坏情形复杂度至少是 $\lceil 3n/2\rceil-2$。

8. 设计一个有效算法，求一列 n 个元素中的次大元素，并确定算法的最坏情形复杂度。

剃须刀排序(或双向冒泡排序)依次比较相邻的元素对，如有逆序就交换它们，交替地从头到尾和从尾到头一遍遍扫描表，直到不需要交换为止。

9. 说明剃须刀排序为表 3，5，1，4，6，2 排序所用的步骤。

10. 用伪码描述剃须刀排序。

11. 求证用比较次数度量，剃须刀排序有 $O(n^2)$ 复杂度。

12. 解释一下为什么剃须刀排序对于接近正确顺序的表的排序是有效的。

13. 求证 $(n\log n+n^2)^3$ 是 $O(n^6)$。

14. 求证 $8^3x+12x+100\log x$ 是 $O(x^3)$。

15. 求 $(x^2+x(\log x)^3)\cdot(2^x+x^3)$ 的大 O 估计。

16. 求 $\sum_{j=1}^{n} j(j+1)$ 的大 O 估计。

*17. 求证 $n!$ 不是 $O(2^n)$。

*18. 求证 n^n 不是 $O(n!)$。

19. 求与5模17同余的四个数。

20. 求证若 a 和 d 为正整数，则有整数 q 和 r，使 $a=dq+r$ 且 $-d/2<r\leqslant d/2$。

*21. 求证若 $ac\equiv bc(\text{mod } m)$，则 $a\equiv b(\text{mod } m/d)$，其中 $d=\gcd(m, c)$。

*22. 在 100_{10}! 的二进制展开中尾部有多少个0?

23. 用欧几里得算法求10 233和33 341的最大公约数。

24. 用欧几里得算法求 gcd(144，233)要做多少次除法?

25. 求 $\gcd(2n+1, 3n+2)$，其中 n 是个正整数。(提示：用欧几里得算法。)

26. a)求证若 a 和 b 为正整数，$a\geqslant b$，则当 $a=b$ 时 $\gcd(a, b)=a$；当 a 和 b 都是偶数时 $\gcd(a, b)=2\gcd(a/2, b/2)$；当 a 为偶数 b 为奇数时 $\gcd(a, b)=\gcd(a/2, b)$；当 a 和 b 都是奇数时 $\gcd(a, b)=\gcd(a-b, b)$。

    b)解释一下如何利用(a)中结论构造一个算法，不用除法，只用比较及二进制展开的减法和移位来求两个正整数的最大公约数。

    c)用这一算法求 gcd(1202，4848)。

27. 求证一个整数能被9整除的充分必要条件是其十进制数字之和能被9整除。

28. a)解释一下为什么 n **div** 7等于 n 天中所含的星期数。

    b)解释一下为什么 n **div** 24等于 n 天中所含的小时数。

一组整数称为是*互素的*，如果其最大公约数是1。

29. 判断下列各组整数是否互素。

    a)8，10，12　　b)12，15，25　　c)15，21，28　　d)21，24，28，32

30. 找一组4个互素的整数，使其中任何两个都不互素。

31. a)假定用函数 $f(p)=(ap+b)$ **mod** 26且 $\gcd(a, 26)=1$ 为信息加密。给出一个可用来解密的函数。

    b)信息加密以后是LJMKGMGMXFQEXMW。如果加密函数是 $f(p)=(7p+10)$ **mod** 26，原信息是什么?

32. 求证同余系统 $x\equiv 2(\text{mod } 6)$ 和 $x\equiv 3(\text{mod } 9)$ 无解。

33. 求出同余系统 $x\equiv 4(\text{mod } 6)$ 和 $x\equiv 13(\text{mod } 15)$ 的所有解。

*34. a)求证同余系统 $x\equiv a_1(\text{mod } m_1)$ 和 $x\equiv a_2(\text{mod } m_2)$ 有解的充分必要条件是 $\gcd(m_1, m_2)\mid a_1-a_2$。

    b)求证(a)中的解模 $\text{lcm}(m_1, m_2)$ 是唯一的。

35. 若 $\mathbf{A}=\begin{bmatrix}0 & 1\\-1 & 0\end{bmatrix}$，求 $\mathbf{A}^n$。

36. 求证若 $\mathbf{A}=c\mathbf{I}$，其中 c 为实数，$\mathbf{I}$ 是 $n\times n$ 单位阵，则对所有 $n\times n$ 矩阵 $\mathbf{B}$ 都有 $\mathbf{AB}=\mathbf{BA}$。

37. 求证若 $\mathbf{A}$ 是2×2矩阵且对任何2×2矩阵 $\mathbf{B}$，都有 $\mathbf{AB}=\mathbf{BA}$，则 $\mathbf{A}=c\mathbf{I}$，其中 c 是实数，$\mathbf{I}$ 是2×2单位阵。

一个 $n\times n$ 矩阵称为*上三角阵*，只要 $i>j$，有 $a_{ij}=0$。

38. 从矩阵乘法的定义出发，设计一个略去自动为0的项的计算两个上三角阵之积的算法。

39. 给出练习38中把两个上三角阵相乘的算法的伪码描述。

40. 练习38中找出的算法用多少元素相乘来计算两个 $n\times n$ 上三角阵之积?

41. 证明若 $\mathbf{A}$ 和 $\mathbf{B}$ 都是可逆矩阵且 $\mathbf{AB}$ 存在，则 $(\mathbf{AB})^{-1}=\mathbf{B}^{-1}\mathbf{A}^{-1}$。

42. 如果 $\mathbf{A}$，$\mathbf{B}$，$\mathbf{C}$，$\mathbf{D}$ 分别是30×10，10×40，40×50和50×30阶的矩阵，计算 $\mathbf{ABCD}$ 的最佳顺序是什么? 假定计算 $p\times q$ 矩阵和 $q\times r$ 矩阵的乘积需要 pqr 次矩阵元素乘法。

43. 令 $\mathbf{A}$ 为 $n\times n$ 矩阵，再令 $\mathbf{O}$ 是元素全为0的 $n\times n$ 矩阵。证明下列各式为真。

    a)$\mathbf{A}\odot\mathbf{O}=\mathbf{O}\odot\mathbf{A}=\mathbf{O}$　　b)$\mathbf{A}\vee\mathbf{O}=\mathbf{O}\vee\mathbf{A}=\mathbf{A}$　　c)$\mathbf{A}\wedge\mathbf{O}=\mathbf{O}\wedge\mathbf{A}=\mathbf{O}$

44. 证明若硬币的面值是 c^0，c^1，…，c^k，其中 k 和 c 是正整数，$c>1$，则贪心算法总是求出用尽可能少硬币的找零钱方案。

45. 设计一个算法猜出一个在1到 2^n-1 之间的数，依次猜出这个数的二进制展开的每一位。

46. 通过依次猜出一个在 1 到 2^n-1 之间的数的二进制展开位来确定这个数，确定就猜的次数而言所需的复杂度。

计算机题目

编写带有下述输入与输出的程序。

1. 已知 n 个整数的表，找出其中的最大整数。
2. 已知 n 个整数的表，找出最大整数在表中的首次和末次出现。
3. 已知 n 个不同整数的表，用线性搜索找出一个整数在表中的位置。
4. 已知 n 个不同整数的有序表，用二分搜索找出一个整数在表中的位置。
5. 已知 n 个整数的表，用冒泡排序为其排序。
6. 已知 n 个整数的表，用插入排序为其排序。
7. 已知一个整数 n，用贪心算法用 25、10、5 和 1 美分硬币找 n 美分零钱。
8. 已知 n 个整数的有序表和一个整数 x，找出用线性搜索和二分搜索确定 x 在表中位置所用的比较次数。
9. 已知一个整数表，确定用冒泡排序和插入排序为其排序所用的比较次数。
10. 已知一个正整数，判断其是否为素数。
11. 已知一段信息，用恺撒密码为其加密；已知一段用恺撒密码加密的信息，为其解密。
12. 已知两个正整数，用欧几里得算法求其最大公约数。
13. 已知两个正整数，求其最小公倍数。

*14. 已知一个正整数，求其素因数分解。

15. 已知一个正整数和一个大于 1 的正整数 b，求这个正整数的 b 进制展开。
16. 已知正整数 a，b 和 $m>1$，求 $a^b \mathbf{mod}\ m$。
17. 已知一个正整数，求其康托展开。（参看 2.5 节练习 44 前的说明。）
18. 已知正整数 n，模数 m，乘数 a，增量 c 和种子 x_0，满足 $0\leqslant a<m$，$0\leqslant c<m$，$0\leqslant x_0<m$，用线性同余发生器 $x_{n+1}=(ax_n+c)\ \mathbf{mod}\ m$ 生成一列 n 个伪随机数。
19. 已知正整数 a 和 b，求整数 s 和 t，使 $sa+tb=\gcd(a, b)$。
20. 已知 n 个以两两互素的素数为模的线性同余式，求这些同余式以这些模数之积为模的公共解。
21. 已知一个 $m\times k$ 矩阵 $\mathbf{A}$ 和一个 $k\times n$ 矩阵 $\mathbf{B}$，求 $\mathbf{AB}$。
22. 已知一个方阵 $\mathbf{A}$ 和一个正整数 n，求 $\mathbf{A}^n$。
23. 已知一个方阵，判断其是否对称。
24. 已知一个 $n_1\times n_2$ 矩阵 $\mathbf{A}$，一个 $n_2\times n_3$ 矩阵 $\mathbf{B}$，一个 $n_3\times n_4$ 矩阵 $\mathbf{C}$ 和一个 $n_4\times n_5$ 矩阵 $\mathbf{D}$，矩阵元素均为整数，确定计算这些矩阵之积的最有效顺序(用整数乘法和加法的次数来度量)。
25. 已知两个 $m\times n$ 布尔矩阵，求其交和并。
26. 已知一个 $m\times k$ 布尔矩阵 $\mathbf{A}$ 和一个 $k\times n$ 布尔矩阵 $\mathbf{B}$，求 $\mathbf{A}$ 和 $\mathbf{B}$ 的布尔积。
27. 已知一个布尔方阵 $\mathbf{A}$ 和一个正整数 n，求 $\mathbf{A}^{[n]}$。

计算和研究

用计算程序或你已经编写的程序做下面的练习。

1. 已知当 b 和 d 是正整数且 $d\geqslant 2$ 时，n^b 是 $O(d^n)$。对于下列这些成对的值：$b=10$ 和 $d=2$，$b=20$ 和 $d=3$，$b=1000$ 和 $d=7$，分别给出 C 和 k 的值，使得只要 $x>k$，就有 $n^b\leqslant Cd^n$。
2. 用贪心算法计算如何用不同面值硬币对不同 n 值来找零钱，并确定是否用了最少的硬币。能否找出保证贪心算法用尽可能少硬币的条件？
3. 利用一个整数 1，2，…，n 的随机序发生器，求冒泡排序、插入排序、二分插入排序和选择排序为这些整数排序所用的比较次数。
4. 对不超过 100 的每个素数 p，判断 2^p-1 是否为素数。
5. 检验某个范围内的大梅森数 2^p-1 以判断其是否为素数。（可以使用 GIMPS 项目的软件。）

6. 证明对 $0 \leqslant n \leqslant 39$ 的所有整数 n，n^2+n+41 是素数，而当 $n=40$ 时它不是素数。当 n 为正整数时，是否有 n 的整系数且次数大于 0 的多项式总是取素数值?
7. 找出尽可能多的形如 n^2+1 的素数，其中 n 为正整数。现在还不知道是否有无穷多个这样的素数。
8. 找出 10 个不同的各有 100 位的素数。
9. 小于 1 000 000 的素数有多少个? 小于 10 000 000 的呢? 小于 100 000 000 的呢? 对于正整数 x，能否提出一个对小于 x 的素数个数的估计?
10. 求随机选取的 10 个不同的 20 位奇数的素数分解。记录分解每个整数消耗的时间。对 10 个 30 位奇数、40 位奇数等尽可能重复上述作业。
11. 求以 2 为基数的所有伪素数，即使得 $2^{n-1} \equiv 1 \pmod n$ 的所有整数 n，其中 n 不超过 10 000。

写作题目

利用本书以外的资料，针对下面的内容写出短文。

1. 查一下算法一词的历史，描述一下早期作品中这一词的用法。
2. 查找一下巴赫曼(Bachmann)论及大 O 记号的原始文献。解释一下他和其他人如何使用这个记号。
3. 解释一下如何根据排序算法所依的基础原理来对排序算法进行分类。
4. 描述一下基数(radix)排序算法。
5. 描述一下并行算法的含义。说明怎样扩展本书使用的伪码以描述并行算法。
6. 解释一下怎样度量并行算法的复杂度。给出一些例子说明这一概念，并说明并行算法可以比没有并行操作的算法更快完成任务。
7. 描述一下用于检验梅森数是否为素数的卢卡斯-莱默尔(Lucas-Lehmer)检验。讨论一下 GIMPS 项目在用这一检验来寻找梅森素数方面的进展。
8. 解释一下实践中怎样用概率素数性检验来产生几乎肯定是素数的特别大的数。这种检验是否有任何潜在的弊端?
9. 75 年前提出的是否有无穷多个卡米切尔数的问题近日得到了解答。描述一下在证明这种数有无穷多个时涉及的要点。
10. 就复杂度和目前能分解的数的大小而言，总结一下因子分解算法的现状。你认为什么时候分解 200 位的数将是可行的?
11. 描述一下现代计算机中使用的正整数加、减、乘、除算法。
12. 描述一下中国剩余定理(孙子定理)的历史。描述中国和印度著作中提出的一些相关问题以及怎样将中国剩余定理用于这些问题。
13. 什么样的一串数是真正的随机数，而不是伪随机数? 用伪随机数做模拟或试验能观察到什么缺点? 伪随机数有哪些性质是随机数不该有的?
14. 描述一下公钥密码的使用现状。就因子分解算法的现状而言，使用公钥密码的方式安全吗? 用公钥密码加密的信息在将来会不会不安全?
15. 描述一下怎样用公钥密码发送签名的保密信息，使得接收方有相当把握确认这个信息是声称发送信息的人所发送的。
16. 说明怎样用同余来判断任意一天是星期几。
17. 描述几个用于大整数有效乘法的算法。
18. 描述几个用于大矩阵有效乘法的算法。
19. 描述一下拉宾(Rabin)公钥密码系统，解释其如何加密和解密信息，以及为什么适合用作公钥密码系统。

第3章　数学推理、归纳与递归

第1章介绍了推理规则和证明方法。第1章后几节和第2章利用这些证明方法证明了各种结果，但是没有讨论过形成猜想、然后尝试确定这些猜想是否正确的过程。本章简要讨论这个过程，提供从证明方法库选择技术和用这些方法构造反例的指南，还讨论反例在这个过程中所起的作用。

本章讨论序列和字符串的某些重要性质，定义可数集合的概念，证明有理数集合是可数的，用所谓对角化论证的证明技术来证明实数集合不是可数的。

许多数学命题都这样断言：对所有正整数来说，某个性质为真。这种命题的例子有：对每个正整数 n 来说，$n! \leqslant n^n$，n^3-n 被3除尽，前 n 个正整数之和是 $n(n+1)/2$ 等。本章和本书的一个主要目标是让学生彻底理解用来证明这类结果的数学归纳法。

第1章和第2章以明确的方式定义了集合和函数，即通过列举集合元素或给出刻画集合元素的某种性质来描述集合；对函数值则给出公式。基于数学归纳法，有另一种重要方式来定义这些对象。要定义函数，就规定某些初始项，给出从已知值来求后续值的规则；定义集合是通过列举某些集合元素，给出从集合中已知元素来构造其他元素的规则。这样的定义称为*递归定义*，在离散数学和计算机科学中大量使用。一旦递归定义了集合，就用所谓的结构归纳法来证明关于这个集合的结果。

解题过程需要保证问题总能够得到正确的解决。仅仅测试出一组输入值结果正确，并不说明这个过程总是正确地工作。只有证明过程总是产生正确结果，才保证了这个过程的正确性。本章最后一节有对程序验证技术的介绍，这是验证过程正确的形式化技术。程序验证是以机械方式证明程序正确的现行尝试的基础。

3.1　证明策略

3.1.1　引言

1.5节介绍了某些最重要的证明方法，说明了如何使用每种方法。1.6节至1.8节和第2章利用这些方法证明了许多不同的定理，但只粗略地讨论了在构造证明背后隐藏的策略。这种策略包括选择证明方法，然后基于这种方法，逐步成功地构造论证。本节将从艺术性与科学性的角度对证明加以研究，给出如何找出定理证明的建议。由于拥有一整套的证明方法，所以更仔细地研究一下如何构造实际的证明。具体地说，描述某些证明技巧，包括如何通过进行回溯、修改已有证明、利用分情形证明法等来进行证明。

数学家的工作就是形成猜想，并试图证明或推翻这些猜想。这里简要描述这个过程，给出关于构造证明和形成猜想的有用建议，除此之外还给出关于反例的某些建议，描述反例在解决某些长期未解决问题中所起的作用，而且讨论一些仍然悬而未决的有趣问题。

最后，我们给出计算机科学中一个极端重要结果的归谬证明，证明不存在这样的过程，当给定计算机程序和输入时，该过程判定输入是否能让程序最终停止。

3.1.2　证明策略

发现证明是挑战性的工作。当面对待证命题时，应该先用定义替换术语，再仔细分析假设和

结论的含义。这样做之后，用一种可用的证明方法去试图证明结果。一般情况下，如果命题是蕴含式，就应该首先尝试直接证明；如果这样不行，就尝试间接证明。如果这些方法都不行，就尝试归谬证明。

前推与后推 无论选择什么证明方法，都需要一个起点。蕴含式的直接证明就从假设开始。利用这些假设以及公理和已知定理，用导向结论的一系列步骤来构造证明。这类推理称为前推，是用来证明相对简单结果的一类最常见推理。同样，要开始间接证明，就从假设的否定开始，用一系列步骤来得出前提的否定。

但是，前推常常难以用来证明更复杂的结果，因为得出想要的结论所需要的推理可能相当复杂。在这种情况下，使用后推可能会有帮助。为了后推证明命题 q，就寻找能证明具有性质 $p\to q$的命题 p。（注意，寻找能证明使得 $q\to r$ 的命题 r 不会有帮助，因为从 $q\to r$ 和 r 得出 q 为真，这种推论是错误的。）例1和例2解释后推。

Extra Examples **例 1** 给定两个正实数 a 和 b，其算术均值是 $(a+b)/2$，其几何均值是 $\sqrt{ab}$。当比较成对不同正实数的算术和几何均值时，发现算术均值总是大于几何均值。（例如，当 $a=4$ 和 $b=6$ 时，有 $5=(4+6)/2>\sqrt{4\cdot 6}=\sqrt{24}$。）能否证明这个不等式恒为真？

解 要证明当 a 和 b 是不同正实数时 $(a+b)/2>\sqrt{ab}$，可进行回溯，构造一系列等价的不等式（请读者来证明每对相邻的不等式都是等价的）。这些等价的不等式是：

$$(a+b)/2>\sqrt{ab} \quad (a+b)^2/4>ab$$
$$(a+b)^2>4ab \quad a^2+2ab+b^2>4ab$$
$$a^2-2ab+b^2>0 \quad (a-b)^2>0$$

由于当 $a\neq b$ 时 $(a-b)^2>0$，所以最后一个不等式为真。由于所有这些不等式都等价，所以当 $a\neq b$ 时 $(a+b)/2>\sqrt{ab}$。一旦构造了这个证明，就能轻而易举地构造使用前推的证明。具体证明留给读者。■

例 2 假定甲乙两人玩游戏，轮流从最初有15块石头的堆中每次取1、2或3块石头。取最后一块石头的人赢得游戏。证明：无论乙如何取，先取的甲都能赢得游戏。

解 为了证明甲总能赢得游戏，进行回溯。在最后一步，如果石头堆中剩下1、2或3块石头，甲就能获胜。如果乙不得不从有4块石头的堆中取石头，就迫使乙留下1、2或3块石头。因此，在倒数第二步甲应该给乙留下4块石头。当乙不得不从8块石头的堆中取石头时，就会留下5、6或7块石头，甲就能留下4块石头。因此，在倒数第三步甲应该给乙留下8块石头。这意味着当轮到甲取时还有9、10或11块石头。同样，当甲走第一步时，应该留下12块石头。把这个论证倒过来就证明了无论乙如何取，甲总是有石头取，所以甲赢得游戏。这些步骤依次给乙留下12、8和4块石头。■

分情形帮助证明 当不能一下子考虑证明的所有情形时，应该考虑分情形证明（在1.5节介绍）。什么时候应该使用这种证明？一般情况下，当没有明显方式可以开始证明，而每种情形中的额外信息又能帮助推动证明时，就寻求分情形证明。例如，为了证明关于整数的结果，分别考虑奇数和偶数，或者分别考虑正整数和负整数（比如涉及到变量的绝对值时），这可能会有好处。

Extra Examples 同样，需要分别考虑整数能被5整除和不能被5整除的情形，或者对某个整数 k 来说，整数形如 $5k$、$5k+1$、$5k+2$、$5k+3$ 或 $5k+4$ 的5种情形。切记分情形证明中的常见错误：许多人忘记考虑全部情形，导致所谓的证明没有对所有可能值来证明结果。例3解释

需要分情形证明的情况。

例 3 证明：如果 n 是不能被 2 或 3 整除的整数，则 n^2-1 能被 24 整除。

解 为了尝试不用分情形的证明，首先注意 $n^2-1=(n-1)\cdot(n+1)$。可是，似乎没有明显的办法来证明当 n 不能被 2 或 3 整除时，这个数能被 24 整除，除非查看不同的情形。分情形证明似乎是下一个合理的方法，但应该使用哪些情形呢？由于想要证明关于不能被 2 或 3 整除的整数的结果，所以对 $j=0$，1，2，3，4，5 来说，分别查看 n 形如 $6k+j$ 的情形可能会有帮助。(根据整除算法，每个整数都具有这些形式之一。)由于只对不能被 2 或 3 整除的整数感兴趣，就立即排除了对某个整数 k 来说 $n=6k$、$6k+2$、$6k+3$ 或 $6k+4$ 的情形。这样只剩下两种情形，其中 n 形如 $6k+1$ 和 $6k+5$。现在分别考虑这两种情形。

假定对某个整数 k 来说 $n=6k+1$。则 $n^2-1=(n-1)(n+1)=(6k+2)6k=12(3k+1)k$。注意 $(3k+1)k$ 是偶数。当 k 是偶数时，这无疑正确；当 k 是奇数时，$3k+1$ 是偶数，所以 $(3k+1)k$ 是偶数。由于 $(3k+1)k$ 是偶数，所以存在整数 q 使得 $(3k+1)k=2q$，因此 $n^2-1=12(3k+1)k=24q$。所以 24 整除 n^2-1。

其次，假定对某个整数 k 来说 $n=6k+5$。则 $n^2-1=(n-1)(n+1)=(6k+6)(6k+4)=12(k+1)(3k+2)$。注意 $(k+1)(3k+2)$ 是偶数，因为当 k 是偶数时，$3k+2$ 是偶数；而当 k 是奇数时，$k+1$ 是偶数。因此存在整数 q 使得 $(k+1)(3k+2)=2q$，所以 $n^2-1=12\cdot 2q=24q$。结论是 n^2-1 能被 24 整除，证毕。 ■

一种常见推理错误是从例子中得出不正确结论。(例如，参见例 8。)除非考虑到每一种可能的情形，否则只考虑个别例子是不能证明定理的。定理的证明类似于证明计算机程序总是产生所需要的输出。无论测试了多少个输入值，都不能得出结论说这个程序总是产生正确输出，除非测试了所有输入值。但例 4 却说明有时候只考虑几个特殊情形就能证明定理。

例 4 证明：$x^2+3y^2=8$ 没有整数解 x 和 y。

解 可以迅速把这个问题归结为只检查很少几种情形，因为当 $|x|\geq 3$ 时 $x^2>8$，当 $|y|\geq 2$ 时 $3y^2>8$。这样还剩下当 x 取 -2、-1、0、1、2 这些值之一以及 y 取 -1、0、1 这些值之一时的情形。为了处理这些情形，注意 x^2 可能的值是 0、1 和 4，$3y^2$ 可能的值是 0 和 3，x^2 和 $3y^2$ 的和最大可能是 7。因此当 x 和 y 是整数时，$x^2+3y^2=8$ 不可能成立。 ■

修改现存证明 一种寻找用来证明命题的好办法是利用现存证明。通常可以修改现存证明来证明新结果。即使这样做不行，现存证明中所用的某些思想仍然可能会有帮助。由于现存证明提供了新证明的线索，所以应当阅读和理解在学习中遇到的证明。例 5 解释如何修改已有证明来证明新结果。

Extra Examples

例 5 证明：存在无穷多个形如 $4k+3$ 的素数，其中 k 是非负整数。

解 或许能够修改的证明是 2.4 节给出的存在无穷多个素数的证明。回忆一下，这个证明假设只存在有穷多个素数 p_1，p_2，…，p_n 并且构造出数 $p_1p_2\cdots p_n+1$。这个数要么是素数，要么具有与素数 p_1，p_2，…，p_n 中每一个都不同的素因子，于是证明了存在无穷多个素数。因此，假设只存在有穷多个形如 $4k+3$ 的素数，即 q_1，q_2，…，q_n，其中 $q_1=3$，$q_2=7$，等等。

构造什么样的数，这个数不能被这些素数中任何一个整除，却必须被形如 $4k+3$ 的素数整除？也许可以考虑数 $4q_1q_2\cdots q_n+3$。但是，这个数不是素数而是能被 3 整除(因为 $q_1=3$)。换一种方法，考虑数

$$Q=4q_1q_2\cdots q_n-1$$

注意 Q 形如 $4k+3$(其中 $k=q_1q_2\cdots q_n-1$)。如果 Q 是素数，就找到了与所列素数都不同并且具有所需形式的素数。如果 Q 不是素数，则 Q 至少有一个不在 q_1，q_2，q_3，…中的素因子，因为当 Q 除以 q_j 时余数是 q_j-1 并且 $q_j-1\neq0$，所以对于 $j=1$，2，…，n 来说 $q_j\times Q$。由于所有的奇素数要么形如 $4k+1$，要么形如 $4k+3$，而且形如 $4k+1$ 的素数之积也具有这种形式(读者可以验证它)，所以 Q 必定具有与所列素数不同的形如 $4k+3$ 的因子，证毕。刚才设法依据已有的证明，做了一些小改动来证明新结果。 ■

数论中只有高深方法才能证明的著名定理说：只要 a 和 b 是互素的正整数，就存在无穷多个形如 $ak+b$ 的素数。请读者尝试是否能依据例 5 中构造的证明，来证明除 $a=4$ 和 $b=3$ 之外，对其他成对 a、b 来说，存在无穷多个形如 $ak+b$ 的素数。

3.1.3 猜想与证明

一般人教数学时方式比较呆板。数学教材(包括本书大部分)形式地给出定理及其证明。这种表达方式没有传达数学的发现过程。这个过程始于探索概念和例子、形成猜想、试图用证明或反例来论证这些猜想。这就是数学家们日复一日的活动。无论是否相信，教科书中的素材原本是以这种方式发展起来的。

人们在各种事实的基础上形成猜想。检查特殊情形导致猜想，确定可能的模式也导致猜想。改变已知定理的假设和结论也导致可信的猜想。在其他时候，基于直觉或者相信结果成立也形成猜想。猜想一旦已经形成，目标就是证明或反驳这个猜想。当数学家们相信猜想可能为真时，就尝试寻找证明。如果找不到证明，就寻找反例。当找不到反例时，就换个方向再次试图证明猜想。尽管许多猜想很快就解决了，但少数猜想几百年都没有解决，导致发展出新的数学分支。例 6 和例 7 解释形成猜想和尝试证明的过程。

例 6 第 2 章提到过已知的最大素数是梅森素数。回忆一下，梅森素数是形如 2^p-1 的素数，其中 p 是素数；即梅森素数是 2 的素数幂减 1。是否还有特殊形式 a^n-1 的其他素数？其中 a 和 n 是正整数。经过某些数值探索(比如注意 $2^6-1=63$、$2^8-1=255$、$3^4-1=80$ 以及 $4^5-1=1023$ 都不是素数)，在发现除了梅森素数以外没有这种形式的其他素数之后，猜想当 $a>2$ 或者当 $a=2$ 且 n 是合数时 a^n-1 是合数。能否证明这个猜想？

如果当 $a>2$ 或者当 $a=2$ 且 n 是合数时找到了 a^n-1 的真因子，就证明了这个猜想。一个有用的方法是利用因式分解 $x^n-1=(x-1)(x^{n-1}+x^{n-2}+\cdots+x+1)$。(回想这个因式分解，如果想不起来，就查阅参考书，比如[Zw02]，这本书列出了特种多项式的因式分解。)在这个因式分解中令 $x=a$，得出 $a-1$ 是 a^n-1 的因子。当 $a=2$ 时有 $a-1=2-1=1$，所以这个因式分解没有产生 a^n-1 的真因子。但是当 $a>2$ 时，因子 $a-1$ 满足 $1<a-1<a^n-1$，说明 a^n-1 不是素数。

还剩下当 $a=2$ 时的情形。当 n 不是素数时，存在正整数 r 和 s，满足 $1<r<n$ 和 $1<s<n$，使得 $n=rs$。利用因子分解 $x^n-1=x^{rs}-1=(x^r-1)(x^{r(s-1)}+\cdots+x^r+1)$，令 $x=a$ 说明 a^r-1 是 a^n-1 的因子。当 $a^r>2$ 时，a^r-1 是大于 1 的正整数。这样当 n 是合数时就产生出 2^n-1 的非平凡因子。 ■

总而言之，下面给出例 6 猜想的简短直接证明。采用分情形证明，因为需要分开考虑 $a>2$ 时与 $a=2$ 时的情形。

定理 1 当 $a>2$ 或者当 $a=2$ 且 n 是合数时，整数 a^n-1 是合数。

证 采用分情形证明。

情形一：当 $a>2$ 时，整数 $a-1$ 是 a^n-1 的因子，因为 $a^n-1=(a-1)(a^{n-1}+a^{n-2}+\cdots+a+1)$ 并且 $1<a-1<a^n-1$。因此 a^n-1 是合数。

情形二：当 $a=2$ 且 n 是合数时，存在整数 r 和 s 满足 $1<r\leqslant s<n$ 使得 $n=rs$。在此情形下，2^r-1 是 2^n-1 的因子，因为 $2^n-1=(2^r-1)(2^{r(s-1)}+\cdots+2^r+1)$ 并且 $1<2^r-1<2^n-1$。这说明 2^n-1 是合数。 □

例 7 能轻而易举地找到一长串相邻合数，比如 24，25，26，27，28 和 90，91，92，93，94，95，96。从这样的证据形成猜想：给定正整数 n，存在 n 个相邻的正合数。如何证明这个猜想？

解 证明这个猜想的一种可能方法是回推。这样推理：如果存在整数 X 使得 X，$X+1$，$X+2$，…，$X+(n-1)$ 都是合数，就证明了结果。（这正是这 n 个整数相邻的含义。）给定 n，能否找到整数 X 使得每个整数 $X+j(j=0, 1, \cdots, n-1)$ 都有非平凡因子？一种可能性是令 $X=n!$。于是 $X=n!$ 是合数，对 $j=2, 3, \cdots, n-1$ 来说，直接看出每个整数 $X+j=n!+j$ 是合数，因为 j 整除 $n!+j$。但不知道 $X+1=n!+1$ 是否为合数。

为了避免这个问题，换一种方法取 $X=(n+1)!+1$。可看到 n 个连续整数 $X+1$，$X+2$，…，$X+n$，即整数

$$(n+1)!+2,(n+1)!+3,\cdots,(n+1)!+(n+1),$$

这些都是合数，因为对 $j=2, 3, \cdots, n+1$ 来说，j 整除 $n!+j$。因此找到了 n 个相邻的合数。改写证明以避免回推的过程，这里就不再详述了。 ■

3.1.4 猜想与反例

有些猜想是似是而非的。要揭示猜想是不正确的，就需要寻找反例。例 8 和例 9 解释这个过程。

Extra Examples **例 8** 对所有正整数 n 来说，函数 $f(n)$ 都是素数。这样的函数 $f(n)$ 很有用，能找到在密码学和其他应用中使用的大素数。一些数学家几百年前就开始检查不同的多项式函数以期找出 $f(n)$。经过许多计算之后发现了多项式 $f(n)=n^2-n+41$。这个多项式具有有趣的性质：对所有不超过 40 的正整数 n 来说，$f(n)$ 是素数。（$f(1)=41$，$f(2)=43$，$f(3)=47$，$f(4)=53$，等等。）这导致猜想：对所有正整数 n 来说，$f(n)$ 是素数。能否验证这个猜想？

解 也许并不奇怪，这个猜想竟然是假的。找个正整数 n 使得 $f(n)$ 是合数并不难，因为 $f(41)=41^2-41+41=41^2$。由于对满足 $1\leqslant n\leqslant 40$ 的所有正整数 n 来说，$f(n)=n^2-n+41$ 是素数，也许会受启发去寻找不同的多项式，它满足性质：对所有正整数 n 来说，$f(n)$ 是素数。但是不存在这样的多项式。可以证明：对每个整系数多项式 $f(n)$，都存在正整数 y，使得 $f(y)$ 是合数（参见练习 37）。 ■

例 9 说明了有的反例大得惊人。重申即使大量数值计算也不能证明定理。

例 9 在 18 世纪，伟大的瑞士数学家列昂哈德·欧拉猜想：对每一个满足 $n\geqslant 3$ 的整数 n 来说，$n-1$ 个正整数的 n 次幂之和不可能是一个 n 次幂。将近 200 年来找到的数值证据都支持这个猜想。但无法找到任何证明，即使对单个的 n 值也不行。然后在 1966 年，L·J·兰德（L. J. Lander）和 T·R·帕金（T. R. Parkin）对 $n=5$ 找到了小得惊人的反例。他们证明了 $27^5+84^5+110^5+133^5=144^5$。

这个反例留下了尚未解决的可能性：对满足 $n\geqslant 3$ 的其他 n 值来说猜想为真。但在 1988 年，诺姆·埃尔基斯（Noam Elkies）对 $n=4$ 找到了反例。所有已知反例都远远大于对 $n=5$ 找到的反例。对 $n=4$，已知最小的反例是 $95\,800^4+217\,519^4+414\,560^4=422\,481^4$（最小的意思是指计算 4 次幂的各个整数之和最小）。满足 $n\geqslant 6$ 的其他情形都没有解决。对 $n=6$ 的情形难以寻找反例，这可能使某人成名，但也有可能（尽管许多人认为不太可能）对 $n=6$ 来说猜想为真。 ■

未解决问题的作用 许多数学进展是与尝试解决著名未解决问题密不可分的。例如，几乎300年都没有解决的一个著名问题导致了整个数论分支的发展。这个问题问：费马大定理是否为真。

定理2 只要 n 是满足 $n>2$ 的整数，方程

$$x^n + y^n = z^n$$

就没有满足 $xyz\neq0$ 的整数解 x、y 和 z。

注意 方程 $x^2+y^2=z^2$ 有无穷多个整数解 x、y 和 z。这些解称为毕达哥拉斯三元组㊀，对应于具有整数边长的直角三角形的边长。参见练习10。

这个问题有一段很有意思的典故。在17世纪，费马在一本丢番图的著作的空白处匆匆写道，他有了"巧妙的证明"：当 n 是大于2的整数时，$x^n+y^n=z^n$ 没有非平凡的整数解。但他从来没有发表过这个证明(费马几乎没有发表过任何东西)，在他死后留下的文章中也找不到任何证明。数学家们花了300年寻找这个证明却没有成功，尽管许多人相信能找到相对简单的证明。(特殊情形的证明，比如欧拉的当 $n=3$ 时情形的证明和费马本人的 $n=4$ 情形的证明，已经找到。)年复一年，少数有声望的数学家认为他们证明了这个定理。在19世纪，这些失败的尝试之一导致了所谓代数数论的数论分支的发展。直到20世纪90年代，当安德鲁·怀尔斯(Adrew Wiles)采用从所谓椭圆曲线论的深奥数论领域中最近发展的思想来证明费马大定理时，才找到了几百页长的高等数学的正确证明。从公共电视台Nova系列的节目中可以看出，利用这个强有力的理论来寻找费马大定理的证明，怀尔斯找了将近10年！(感兴趣的读者可以查阅[Ro99]来了解关于费马大定理的更多信息和关于这个问题及其解决的其他参考资料。)

许多著名问题还在等待聪明人的最终解决。例10至例13描述这些悬而未决问题中最容易理解和比较著名的几个。

例10 哥德巴赫㊀猜想 在1742年，克里西安·哥德巴赫(Christian Goldbach)在给列昂哈德·欧拉(Leonhard Euler)的信中猜想：每个奇数 $n(n>5)$ 都是3个素数之和。欧拉回答说这个猜想等价于每个偶数 $n(n>2)$ 都是两个素数之和的猜想(参见练习39)。每个偶数 $n(n>2)$ 都是两个素数之和的猜想称为*哥德巴赫猜想*。对小的偶数能验证这个猜想。例如，$4=2+2$，$6=3+3$，$8=5+3$，$10=7+3$，$12=7+5$，等等。在计算机出现之前，对于直到几百万的数字，手工计算都验证了哥德巴赫猜想。借助计算机，对非常大的数字做了验证。到2002年中期，对于直到 $4\cdot10^{14}$ 的所有正偶数都验证了猜想。

尽管没有找到哥德巴赫猜想的证明，大多数数学家还是相信猜想为真。采用复杂的解析数论方法已经证明了若干定理，证明了比哥德巴赫猜想更弱的结果，其中有：每个大于2的偶数都是至多6个素数之和(O. Ramaré在1995年证明)，以及：每个足够大的正整数都是素数与另一数之和，这个数是素数或两个素数之积(陈景润在1966年证明)。或许哥德巴赫猜想在不远的将来就会被解决。 ■

例11 有许多猜想断言：存在无穷多个具有特定形式的素数。例如，有这样的猜想：存在无穷多个梅森素数。(回忆一下2.4节，这些素数是形如 2^p-1 的素数，其中 p

---

㊀ 也叫做勾股数组。——译者注

㊀ **克里西安·哥德巴赫(Christian Goldbach，1690—1764)**诞生在普鲁士的哥尼斯堡，这个城市因著名的七桥问题而闻名(8.5节研究该问题)。他在1725年成为圣彼得堡科学院的数学教授。哥德巴赫在1728年到莫斯科当沙皇儿子的家庭教师。他在1742年成为俄国外交部的职员，进入政界。哥德巴赫最出名的，是他与包括欧拉和伯努利在内的杰出数学家们的通信、他在数论里的著名猜想以及对数学分析的若干贡献。

是素数。)另一个这种猜想是：存在无穷多个形如 n^2+1 的素数，其中 n 是正整数。例如，$5=2^2+1$，$17=4^2+1$，$37=6^2+1$，等等。目前已知的最好结果是：存在无穷多个正整数 n，使得 n^2+1 是素数或至多两个素数之积(Henryk Iwaniec 在 1973 年证明)。 ■

例 12　孪生素数猜想　孪生素数是相差 2 的素数，比如 3 和 5、5 和 7、11 和 13、17 和 19 以及 4967 和 4969 等。孪生素数猜想断言：存在无穷多个孪生素数。关于孪生素数证明的最有力结果是：存在无穷多对 p 和 $p+2$，其中 p 是素数并且 $p+2$ 是素数或两个素数之积(陈景润在 1966 年证明)。到 2002 年中期时，孪生素数的世界记录包含数字 $318\,032\,361\cdot 2^{107\,001}\pm 1$，一个具有 32 220 位的数字。 ■

例 13　3x+1 猜想　设 $f(x)$是把偶数 x 映射到 $x/2$、把奇数 x 映射到 $3x+1$ 的函数。有个著名的猜想(有时称为 $3x+1$ 猜想)说：对所有正整数 x 来说，当反复地应用函数 f 时，最终会得到整数 1。例如，从$x=13$ 开始，发现 $f(13)=3\cdot 13+1=40$，$f(40)=40/2=20$，$f(20)=20/2=10$，$f(10)=10/2=5$，$f(5)=3\cdot 5+1=16$，$f(16)=8$，$f(8)=4$，$f(4)=2$，$f(2)=1$。对于直到 $5.6\cdot 10^{13}$ 的所有整数都验证了 $3x+1$ 猜想。

$3x+1$ 猜想具有有趣的历史，从 20 世纪 50 年代以来就吸引了数学家们的注意力。这个猜想被多次提出，具有许多其他名称，包括：Collatz 问题、Hasse 算法、Ulam 问题、Syracuse 问题以及 Kakutani 问题等。许多数学家抛开原有工作来花时间解决这个猜想。这引起一则笑话，说这个问题是旨在减缓美国数学研究的阴谋的一部分。参见 Jeffrey Lagaris 的文章[La85]来了解对这个问题有意思的讨论以及试图解决这个问题的数学家们所发现的结果。 ■

3.1.5　停机问题

现在描述计算机科学里一个最著名定理的证明，证明存在不能用任何过程来解决的问题，即 2.3 节提到过的，证明存在不可解的问题。研究的这个问题是*停机问题*，问是否存在这样做的过程：以计算机程序和程序的输入作为输入，确定当程序依据这个输入运行时是否最终停止。可以以此帮助测试程序是否进入死循环，当编写和调试程序时肯定会有帮助。但在 1936 年，阿兰·图灵(Alan Turing)证明了不存在这样的过程(参见 11.4 节他的生平介绍)。

在给出停机问题是不可解的证明之前先要注意，运行程序并观察程序做什么，仅仅如此并不能确定当程序在给定输入上运行时是否终止。如果程序停止，就得出了答案，但如果经过任何固定长度的时间之后程序仍在运行，就不知道究竟是永不停机，还是没有等到足够长的停机时间。毕竟运算 10 亿年之后才停止的程序是不难设计出来的。

下面描述停机问题是不可解的图灵证明；这是归谬证明。(读者应当注意，本节的证明不是十分严格，因为没有明确定义什么是过程。要补救这一点，就需要图灵机概念，11.5 节介绍这个概念。)

证　假定停机问题有解，即存在所谓 $H(P, I)$的过程。过程 $H(P, I)$有两个输入，一个是程序 P，另一个是程序 P 的输入 I。如果 H 确定当给定 I 作为输入时 P 停止，则 $H(P, I)$产生字符串“停机”作为输出。否则，$H(P, I)$产生字符串“死循环”作为输出。现在导出矛盾。

过程在编码时表示成字符串；把这个字符串解释成比特序列。这意味着程序本身可用作数据。因此可认为程序是另一个程序的输入，甚至是自身的输入。因此，H 可用程序 P 作为两个输入，即程序和这个程序的输入。H 应当能够确定当 P 给定自身的拷贝作为输入时是否停机。

要证明不存在解决停机问题的过程 H，就构造简单过程 $K(P)$，$K(P)$利用 $H(P, P)$的输出来这样工作：如果 $H(P, P)$的输出是“死循环”，这意味着当 P 给定自身拷贝作为输入时死循

环，则 $K(P)$停机；如果 $H(P, P)$的输出是“停机”，这意味着当 P 给定自身拷贝作为输入时停机，则 $K(P)$死循环。即 $K(P)$做与 $H(P, P)$所规定的输出相反的动作(参见图 3-1)。

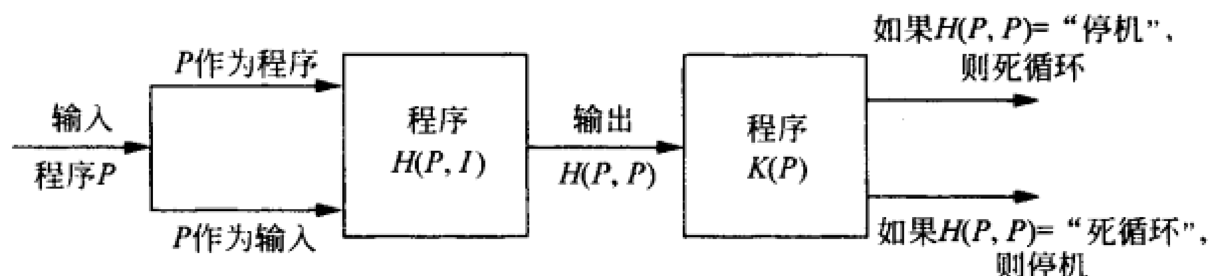


图 3-1

现在假定把 K 作为 K 的输入。注意，如果 $H(K, K)$的输出是“死循环”，则根据 K 的定义可知 $K(K)$停机；否则，如果 $H(K, K)$的输出是“停机”，则根据 K 的定义可知 $K(K)$死循环，与 H 所说的相悖。在两种情形下都有矛盾。

因此 H 不能总是给出正确答案。所以不存在解决停机问题的过程。 □

3.1.6 其他证明方法

第 1 章介绍了证明中使用的基本方法，本节描述如何利用这些方法来证明各种结果。但除了这些讨论过的方法之外还有许多重要的证明方法，本书稍后介绍若干这种方法。具体地说，3.3 节讨论数学归纳法，即证明形如 $\forall nP(n)$的命题的极为有用的方法，其中论域是正整数集合。3.4 节介绍结构归纳法，用来证明关于递归定义的集合的结果。3.2 节使用康托对角化方法，用来证明关于无穷集大小的结果。最后，第 4 章介绍组合证明的概念，用来证明采用计数论证的结果。读者应当注意已经有相关书籍专门描述本节讨论的内容，包括乔治·波利亚(George Polya)的许多优秀著作([Po61]，[Po71]，[Po90])。

练习

1. 证明：任何三个相邻整数之积能被 6 整除。
2. 证明：如果 n 是奇正整数，则 $n^2 \equiv 1 \pmod 8$。
3. 证明：如果 n 是奇正整数，则 $n^4 \equiv 1 \pmod{16}$。
4. 证明：方程 $2x^2+5y^2=14$ 没有整数解 x 和 y。
5. 证明：方程 $x^4+y^4=625$ 没有整数解 x 和 y。
6. 证明：没有整数 x 和 y 使得 $x^2-4y=3$。
7. 证明：没有整数 x 和 y 使得 $3x^2-8y=1$。
8. 证明：方程 $x^2-5y^2=2$ 没有整数解 x 和 y。(提示：考虑这个方程模 5。)
9. 证明：方程 $x^4-16y^4=3$ 没有整数解 x 和 y。(提示：考虑这个方程模 16 并利用练习 3，观察所得到的同余式。)
10. 证明：方程 $x^2+y^2=z^2$ 有无穷多个正整数解 x、y 和 z。(提示：令 $x=m^2-n^2$、$y=2mn$ 以及 $z=m^2+n^2$，其中 m 和 n 是整数。)
11. 两个整数 a 和 b 的调和均值是$\frac{2ab}{a+b}$。通过计算不同成对正整数的调和均值和几何均值，形成关于这两种均值相对大小的猜想并证明之。
12. 两个整数 a 和 b 的平方均值是$\sqrt{\frac{a^2+b^2}{2}}$。通过计算不同成对正整数的算术均值和平方均值，形成关于这

两种均值相对大小的猜想并证明之。

*13. 在黑板上写数字 1，2，…，$2n$，其中 n 是奇数，从中挑出两个数 j 和 k，在黑板上写 $|j-k|$ 并擦掉 j 和 k。继续这个过程，直到黑板上只写一个整数为止。证明：这个整数必为奇数。

*14. 假设 5 个 1 和 4 个 0 绕圆周排列。在任何两个相同的位之间插入 1，在任何两个不同的位之间插入 0，以产生 9 个新的位。然后删除原来的 9 个位。证明：当反复进行这个过程时，永远不能得到 9 个 0。（提示：进行回溯，假设真的以 9 个 0 结束。）

*15. 证明或反驳：每当 n 是正整数时，$n^2-79n+1601$ 是素数。

16. 证明或反驳：对所有非负整数 n 来说，2^n+1 是素数。

17. 证明或反驳：$\sqrt[3]{3}$是无理数。

*18. 证明：若 n 是正整数但不是完全平方数，则$\sqrt{n}$是无理数。

*19. 证明或反驳：不存在有理数 x，使得 $x^3+x+1=0$。

20. 证明：当不能被 5 整除的整数的平方除以 5 时，余数是 1 或 4。（提示：使用分情形证明。）

*21. 设 p 是素数。证明：$a^2\equiv b^2 \pmod p$，当且仅当 $a\equiv(b \pmod p)$或 $a\equiv -b \pmod p$。

22. 证明或反驳：每当 n 是大于 1 的正整数时，n^2-1 是合数。

23. 证明或反驳：每当 m 是正整数时，对所有整数 a 和 b 来说，$(a \bmod m)(b \bmod m)=ab \bmod m$。

24. 证明或反驳：每当 m 是正整数时，对所有整数 a 和 b 来说，$a \bmod m+b \bmod m=(a+b) \bmod m$。

25. 证明或反驳：存在 3 个相邻正奇数都是素数，即存在形如 p、$p+2$ 和 $p+4$ 的奇素数。

26. 证明或反驳：给定正整数 n，存在 n 个相邻正奇数都是素数。

27. 给出下列命题的构造性证明："对每个正整数 n 来说，都存在能被 n 个以上素数整除的整数"。

28. 找出下列命题的反例："对每个素数 n 来说，$n+2$ 是素数"。

29. 修改例 5 给出的证明，证明：存在无穷多个形如 $6k+5$ 的素数。

30. 当试图利用存在无穷多个素数的证明去证明存在无穷多个形如 $4k+1$ 的素数时，会出现什么问题？

*31. 证明：如果 x 是实数且 n 是正整数，则

$$\left\lfloor\frac{x}{3}\right\rfloor+\left\lfloor\frac{x+n}{3}\right\rfloor+\left\lfloor\frac{x+2n}{3}\right\rfloor$$
$$=\gcd(3,n)\left\lfloor\frac{x}{\gcd(3,n)}\right\rfloor+(n-1)+\frac{\gcd(3,n)-1}{2}。$$

32. 使用分情形证明来证明对所有整数 n 来说：$\lfloor n/2\rfloor\lceil n/2\rceil=\lfloor n^2/4\rfloor$。

*33. 设 $S=x_1y_1+x_2y_2+\cdots+x_ny_n$，其中 x_1，x_2，…，x_n 和 y_1，y_2，…，y_n 是排序的两个不同正实数序列，各自有 n 个元素。

a)证明：在这两个序列的所有排序当中，当两个序列都排序(使得每个序列中的元素都以非降序排列)时，S 取最大值。

b)证明：在这两个序列的所有排序当中，当一个序列排成非降顺序，另一个序列排成非升顺序时，S 取最小值。

34. 证明：若 m 和 n 都是正整数而 x 是正实数，则

$$\left\lfloor\frac{\lfloor x\rfloor+n}{m}\right\rfloor=\left\lfloor\frac{x+n}{m}\right\rfloor$$

*35. 证明：若 m 是正整数而 x 是正实数，则

$$\lfloor mx\rfloor=\lfloor x\rfloor+\left\lfloor x+\frac{1}{m}\right\rfloor+\left\lfloor x+\frac{2}{m}\right\rfloor+\cdots+\left\lfloor x+\frac{m-1}{m}\right\rfloor$$

**36. 证明：若 a 和 b 都是正无理数使得 $1/a+1/b=1$，则每个正整数都唯一表示成$\lfloor ka\rfloor$或$\lfloor kb\rfloor$，其中 k 是某个正整数。

*37. 证明：若 $f(x)$是整系数非常值多项式，则存在整数 y 使得 $f(y)$是合数。（提示：假设 $f(x_0)=p$ 是素数。证明对所有整数 k 来说，p 整除$f(x_0+k_p)$。得出与下述事实的矛盾：n 次多项式最多 n 次取得同一个值，其中 n 大于 1。）

38. 证明：若 n 是正整数使得其因子和是 $n+1$，则 n 是素数。
39. 证明：哥德巴赫猜想(这个猜想说每个大于 2 的偶数都是两个素数之和)等价于每个大于 5 的整数都是三个素数之和。
40. 所谓的德波里格耐克(de Polignac)猜想说每个正偶数都可以用无穷多种方式写成两个相邻素数之差。证明：德波里格耐克猜想蕴含孪生素数猜想。
41. 对下列这些整数验证 $3x+1$ 猜想：

    a)6　　b)7　　c)17　　d)21
42. 对下列这些整数验证 $3x+1$ 猜想：

    a)16　　b)11　　c)35　　d)113

完满数是等于其真因子(即除自身外的因子)和的正整数。

43. 证明：6、28 和 496 都是完满数。
44. 证明：如果 $\gcd(s, t)=1$，其中 s 和 t 是正整数，则 st 的因子和是 s 的因子和与 t 的因子和之积。
45. 证明：当 2^p-1 是梅森素数时，整数 $2^{p-1}(2^p-1)$ 是完满数。

**46. 证明：若 n 是完满数，则 $n=2^p(2^p-1)$，其中 2^p-1 是梅森素数。

47. 证明或反驳：如果有 8 加仑的一罐水，并有容量分别为 5 加仑和 3 加仑的两个空罐，那么通过反复把一个罐中的部分或全部水注入另一个罐，就可以量出 4 加仑的水。

*48. 证明或反驳：若 n 是正整数，则 $\lfloor\sqrt{n}+\sqrt{n+1}\rfloor=\lfloor\sqrt{4n+2}\rfloor$。

49. 证明：确定带有给定输入的程序是否会打印数字 1 的问题是不可解的。
50. 证明：判定带有具体输入的具体程序是否停机的问题是可解的。
51. 证明：下列问题是可解的。给定两个程序及其输入并知道其中恰有一个停机，确定哪一个停机。

3.2 序列与求和

3.2.1 引言

序列用来表示元素的有序表。在离散数学中用许多方式使用序列。第 6 章将介绍用序列表示某些计数问题的解。在计算机科学中序列也是重要的数据结构。本节包括对函数概念的回顾以及用来表示序列及序列项之和的记号。

当可以列举无穷集合的元素时，这个集合称为可数的。本节末尾讨论可数和不可数集合，证明有理数集合是可数的而实数集合是不可数的。

3.2.2 序列

序列是用来表示有序表的离散结构。

定义 1 序列是从整数集合的子集(通常是{0, 1, 2, …}或{1, 2, 3, …})到集合 S 的函数。用记号 a_n 表示整数 n 的像。称 a_n 为序列的一个项。

用记号 $\{a_n\}$ 描述序列。(注意 a_n 表示序列 $\{a_n\}$ 的个别项。还要注意序列记号 $\{a_n\}$ 与集合记号有冲突，但使用这个记号的上下文总是分得清楚何时在处理序列而何时在处理集合。还要注意字母 a 的选择是任意的。)

描述序列是通过按照下标升序来列举序列项。

例 1 考虑序列 $\{a_n\}$，其中

$$a_n = 1/n$$

从 a_1 开始的这个序列项的表，即

$$a_1, a_2, a_3, a_4, \cdots,$$

开头是：

$$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots$$ ■

定义 2 几何序列是如下形式的序列

$$a, ar, ar^2, \cdots, ar^n$$

其中初项 a 和公比 r 都是实数。

注意 几何序列是指数函数 $f(x)=ar^x$ 的离散对应物。

例 2 满足 $b_n=(-1)^n$、$c_n=2 \cdot 5^n$ 和 $d_n=6 \cdot (1/3)^n$ 的序列$\{b_n\}$、$\{c_n\}$和$\{d_n\}$都是几何序列，初项和公比分别等于-1 和-1、10 和 5 以及 2 和 1/3。表 b_1，b_2，b_3，b_4，…的开头是：

$$-1, 1, -1, 1, \cdots$$

表 c_1，c_2，c_3，c_4，…的开头是：

$$10, 50, 250, 1250, \cdots$$

表 d_1，d_2，d_3，d_4，…的开头是：

$$2, 2/3, 2/9, 2/27, \cdots$$ ■

定义 3 等差序列是如下形式的序列

$$a, a+d, a+2d, \cdots, a+nd$$

其中初项 a 和公差 d 都是实数。

注意 等差序列是线性函数 $f(x)=dx+a$ 的离散对应物。

例 3 满足 $s_n=-1+4n$ 和 $t_n=7-3n$ 的序列$\{s_n\}$和$\{t_n\}$都是等差序列，初项和公差分别等于-1 和 4 以及 7 和-3。从 $n=0$ 这项开始的表 s_0，s_1，s_2，s_3，…的开头是：

$$-1, 3, 7, 11, \cdots$$

表 t_0，t_1，t_2，t_3，…的开头是：

$$7, 4, 1, -2, \cdots$$ ■

在计算机科学中经常使用形如 a_1，a_2，…，a_n 的序列。这些有穷序列也称为串。这个串也记作 $a_1a_2\cdots a_n$。(回忆一下，位串是比特位的有限序列，在 1.1 节介绍)。串 S 的长度是这个串中的项数。空串是没有任何项的串，记作 λ。空串长度为 0。

例 4 串 $abcd$ 是长度为 4 的串。 ■

3.2.3 特殊的整数序列

离散数学中的常见问题是找出构造序列项的公式或一般规则。有时候只知道解决问题的序列的一部分项，目标是确定序列。尽管序列的初始项不能确定整个序列(毕竟存在无穷多个从任何有穷初始项集合开始的序列)，但知道前几项仍有助于形成关于序列本身的合理猜想。一旦形成猜想，就可以尝试验证找到了正确序列。

当试图从初始项推导序列项的可能公式或规则时，尝试一下找出这些项中的模式，再观察一下能否确定如何从前项产生后项。有许多问题可问，但某些较有用的问题是：

- 是否有相同值连续出现?
- 是否给前项加上常量或与序列中位置有关的量就得出后项?

- 是否给前项乘以特定量就得出后项？
- 是否按照某种方式组合前项就得出后项？
- 是否在各项之间存在循环？

Extra Examples **例5** 求具有下列前5项的序列公式：(a)1，1/2，1/4，1/8，1/16 (b)1，3，5，7，9 (c)1，−1，1，−1，1。

解 (a)可以看出分母都是2的幂。满足$a_n=1/2^{n-1}$的序列是候选解序列。这个候选序列是几何序列，满足$a=1$和$r=1/2$。

(b)注意每一项都是对前一项加2而得到的。满足$a_n=2n-1$的序列是候选解序列。这个候选序列是等差序列，满足$a=1$和$d=2$。

(c)各项轮流取值1和−1。满足$a_n=(-1)^{n+1}$的序列是候选解序列。这个候选序列是几何序列，满足$a=1$和$r=-1$。 ■

例6和例7说明如何通过分析序列来发现构造项。

例6 如果前10项是1，2，2，3，3，3，4，4，4，4，则如何产生序列项？

解 注意整数1出现1次，整数2出现2次，整数3出现3次，整数4出现4次。产生这个序列的合理规则是整数n恰好出现n次，所以序列的下5项可能都是5，随后6项可能都是6，等等。这种方式产生的序列是候选解序列。 ■

例7 如果前10项是5，11，17，23，29，35，41，47，53，59，则如何产生序列项？

解 注意这个序列第一项之后的前10项每项都是给前项加6得到的(从相邻项之差为6就看出这一点)。因此从5开始总共加$(n-1)$次6就产生第n项，合理猜测第n项为$5+6(n-1)=6n-1$。(这是个等差序列，满足$a=5$和$d=6$。) ■

另一种求序列项生成规则的有用技术是对比所求的序列项与熟知的整数序列项，比如等差序列项、几何序列项、完全平方、完全立方，等等。表3-1显示若干应当记住的序列的前10项。

表3-1 若干有用序列

第n项	前10项
n^2	1，4，9，16，25，36，49，64，81，100，…
n^3	1，8，27，64，125，216，343，512，729，1000，…
n^4	1，16，81，256，625，1296，2401，4096，6561，10 000，…
2^n	2，4，8，16，32，64，128，256，512，1024，…
3^n	3，9，27，81，243，729，2187，6561，19 683，59 049，…
$n!$	1，2，6，24，120，720，5040，40 320，362 880，3 628 800，…

例8 如果序列$\{a_n\}$的前10项为1，7，25，79，241，727，2185，6559，19 681，59 047，试猜想a_n的简单公式。

解 要解决这个问题，先查看相邻项的差，但没有看出模式。当计算相邻项的比来查看每项是否为前项的倍数时，发现这个比虽然不是常数却接近于3。所以有理由怀疑与3^n有关的公式产生了这个序列的各项。比较这些项与序列$\{3^n\}$的对应项，注意到第n项要比对应的3的幂少2。看到对于$1\leqslant n\leqslant 10$来说$a_n=3^n-2$，因而猜想对所有n来说，这个公式成立。 ■

贯穿本书，可以看到整数序列广泛出现在离散数学中。已经遇到或即将遇到的序列包括：素数序列(第2章)、排序n个离散对象的方式数(第4章)、解决著名的n个盘子汉诺塔智力题需要的步数(第6章)以及在n个月后岛上的兔子数(第6章)等。

整数序列还出现在离散数学以外的十分广泛的专题领域中，包括生物学、工程、化学、物理学以及智力题等。数学家内尔·斯朗[1](Neil Sloane)在过去20年中构造了超过8000个五花八门妙趣横生的整数序列，他还与西蒙·普鲁夫(Simon Plouffe)合作出版了《整数序列百科全书》(*The Encyclopedia of Integer Sequences*)([SlPl95])。经过扩充的序列表出现在网络上，定期地加入新序列。还有通过网络访问的程序，用来从百科全书中查找给定初始项的解序列。

3.2.4 求和

下一步介绍求和记号。首先描述用来表达序列$\{a_n\}$中

$$a_m, a_{m+1}, \cdots, a_n$$

这些项之和的记号。用记号

$$\sum_{j=m}^{n} a_j \text{ 或 } \sum_{j=m}^{n} a_j$$

表示

$$a_m + a_{m+1} + \cdots + a_n$$

此处变量j称为求和下标，字母j是任意选择作为变量的；换句话说，可以使用任何其他字母，比如i或k。或者用记号写成

$$\sum_{j=m}^{n} a_j = \sum_{i=m}^{n} a_i = \sum_{k=m}^{n} a_k$$

此处求和下标依次等于从下限m开始到上限n为止的所有整数。用大写希腊字母Σ表示求和。下面给出求和记号的若干例子。

例 9 表示序列$\{a_n\}$前100项之和，对$n=1, 2, 3, \cdots$，$a_n=1/n$。

解 求和下标下限为1，上限为100。把这个和写成

$$\sum_{j=1}^{100} \frac{1}{j}$$

■

例 10 $\sum_{j=1}^{5} j^2$的值是什么？

解 我们有

$$\sum_{j=1}^{5} j^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2$$

---

[1] **内尔·斯朗(Neil Sloane，生于1939年)**依靠澳大利亚国家电话公司的助学金在墨尔本大学学习数学和电气工程。他在暑期中接触了许多与电话有关的工作，比如树电话杆。毕业后，他设计了澳大利亚最便宜的电话网。1962年他到美国康奈尔大学学习电气工程。他的博士论文是关于现在所谓的神经网络。自1969年开始，他在贝尔实验室任职，从事过许多领域的工作，包括：网络设计、编码理论以及球体打包等。1996年AT&T从贝尔实验室分离出来至今，他一直在那里工作。他最喜欢的一个问题是相切问题(他创造的名字)，问在n维空间可以排列多少个球，使之都与同样大小的中央球相切。(在二维答案是6，因为可放置6个便士使之与中央便士相切。在三维，可放置12个台球使之与中央台球相切。两个恰好在一点接触的台球称为“相切”，这引出了术语“相切问题”和“相切数”)。斯朗与Andrew Odlyzko证明了在8维和24维中最优相切数分别为240和196 560。在1，2，3，8和24维中的相切数是已知的，而在任何其他维中都是未知的。斯朗的著作包括：与John Conway合写的《*Sphere Packing, Lattices and Groups*》第3版，与Jessie Mac Williams合写的《*The Theory of Error-Correcting Codes*》，与Simon Plouffe合写的《*The Encyclopedia of Integer Sequences*》以及与Paul Nick合写的《*The Rock-Climbing Guide to New Jersey Crags*》。最后这本书体现了他对攀岩的兴趣，其中介绍了新泽西州的50多个攀岩场地。

$$= 1 + 4 + 9 + 16 + 25$$
$$= 55$$ ■

例 11 $\sum_{k=4}^{8}(-1)^k$ 的值是什么?

解 我们有

$$\sum_{k=4}^{8}(-1)^k = (-1)^4 + (-1)^5 + (-1)^6 + (-1)^7 + (-1)^8$$
$$= 1 + (-1) + 1 + (-1) + 1 = 1$$ ■

有时候平移一下和中的求和下标会有用处。通常这样做是当两个和需要相加而求和下标却不一致时。当平移求和下标时,重要的是适当地修改对应求和项。下面的例子解释这一点。

例 12 假定有和

$$\sum_{j=1}^{5} j^2$$

但是希望求和下标在 0 和 4 之间(而不是在 1 和 5 之间)依次取值。为此令 $k=j-1$。于是新的求和下标从 0 到 4 依次取值,项 j^2 变成$(k+1)^2$。因此

$$\sum_{j=1}^{5} j^2 = \sum_{k=0}^{4}(k+1)^2$$

容易验证两个和都是 1+4+9+16+25=55。 ■

几何序列各项之和经常出现(这种和称为几何级数)。定理 1 给出几何序列各项之和的公式。

定理 1 若 a 和 r 都是实数且 $r\neq 0$,则

$$\sum_{j=0}^{n} ar^j = \begin{cases} \dfrac{ar^{n+1}-a}{r-1} & 若\ r\neq 1 \\ (n+1)a & 若\ r=1 \end{cases}$$

证 设

$$S = \sum_{j=0}^{n} ar^j$$

要计算 S,先给等式两边同乘 r,然后如下变换得出的和:

$$rS = r\sum_{j=0}^{n} ar^j = \sum_{j=0}^{n} ar^{j+1}$$
$$= \sum_{k=1}^{n+1} ar^k = \sum_{k=0}^{n} ar^k + (ar^{n+1}-a) \quad (令\ k=j+1,平移求和下标即得这个等式)$$
$$= S + (ar^{n+1}-a)$$

从这些等式看出

$$rS = S + (ar^{n+1}-a)$$

解出 S 就证明了若 $r\neq 1$

$$S = \frac{ar^{n+1}-a}{r-1}$$

若 $r=1$,则显然这个和等于$(n+1)a$。 □

例 13 很多情况下要求双重求和(比如在计算机程序中嵌套循环的分析中)。双重求和的例子是

$$\sum_{i=1}^{4}\sum_{j=1}^{3} ij$$

要计算双重求和，先展开内层求和，再继续计算外层求和：

$$\sum_{i=1}^{4}\sum_{j=1}^{3} ij = \sum_{i=1}^{4}(i+2i+3i) = \sum_{i=1}^{4} 6i$$
$$= 6+12+18+24 = 60$$ ■

还用求和记号把函数的所有值相加，或把指标集的各项加起来，其中求和下标依次等于集合中的所有值。换句话说，用

$$\sum_{s\in S} f(s)$$

表示对 S 中所有元素 s 的值 $f(s)$ 求和。

例 14 $\sum_{s\in\{0,2,4\}} s$ 的值是什么？

解 由于 $\sum_{s\in\{0,2,4\}} s$ 表示对集合$\{0,2,4\}$中所有元素 s 的值求和，所以

$$\sum_{s\in\{0,2,4\}} s = 0+2+4 = 6$$ ■

某些求和在离散数学中反复出现。掌握这种求和的一组公式会有用处，所以表 3-2 为常见求和提供袖珍公式表。

表 3-2 若干有用的求和公式

和	闭 形 式
$\sum_{k=0}^{n} ar^k (r\neq 0)$	$\frac{ar^{n+1}-a}{r-1}, r\neq 1$
$\sum_{k=1}^{n} k$	$\frac{n(n+1)}{2}$
$\sum_{k=1}^{n} k^2$	$\frac{n(n+1)(2n+1)}{6}$
$\sum_{k=1}^{n} k^3$	$\frac{n^2(n+1)^2}{4}$
$\sum_{k=0}^{\infty} x^k, \lvert x\rvert<1$	$\frac{1}{1-x}$
$\sum_{k=1}^{\infty} kx^{k-1}, \lvert x\rvert<1$	$\frac{1}{(1-x)^2}$

定理 1 推导了表中的第一个公式。接下来三个公式给出了前 n 个正整数的求和、平方求和以及立方求和。可用许多不同方式推导这三个公式(例如，参见本节末练习 21 和练习 22)。还要注意一旦知道了这些公式中的每一个，就能用数学归纳法(3.3 节的主题)轻而易举地加以证明。表中最后两个公式与无穷级数有关，很快就会讨论到。

例 15 解释了表 3-2 中的公式是如何使用的。

例 15 求 $\sum_{k=50}^{100} k^2$。

解 首先注意由于 $\sum_{k=1}^{100} k^2 = \sum_{k=1}^{49} k^2 + \sum_{k=50}^{100} k^2$，所以有

$$\sum_{k=50}^{100} k^2 = \sum_{k=1}^{100} k^2 - \sum_{k=1}^{49} k^2$$

利用表 3-2 的公式 $\sum_{k=1}^{n} k^2 = n(n+1)(2n+1)/6$，看出

$$\sum_{k=50}^{100} k^2 = \frac{100 \cdot 101 \cdot 201}{6} - \frac{49 \cdot 50 \cdot 99}{6} = 38\ 350 - 40\ 425 = 297\ 925 \quad \blacksquare$$

若干无穷级数 尽管本书中大多数求和都是有穷求和，但在离散数学某些部分中无穷级数却是重要的。例 16 和例 17 中无穷级数的闭形式非常有用。

例 16 （本例要求微积分知识）设 x 是实数满足 $|x|<1$。求 $\sum_{n=0}^{\infty} x^n$。

解 根据定理 1 让 $a=1$ 和 $r=x$ 就看出 $\sum_{n=0}^{k} x^n = \frac{x^{k+1}-1}{x-1}$。由于 $|x|<1$，所以当 k 趋于无穷时 x^{k+1} 趋于 0。所以

$$\sum_{n=0}^{\infty} x^n = \lim_{k\to\infty} \frac{x^{k+1}-1}{x-1} = \frac{-1}{x-1} = \frac{1}{1-x} \quad \blacksquare$$

对已有公式进行微分或积分就产生新的求和公式。

例 17 （本例要求微积分知识）对下列方程两边微分

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

根据例 16 就发现

$$\sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}$$

（根据关于无穷级数的定理，对 $|x|<1$ 这个微分有效。） ■

3.2.5 基数

回忆一下，1.6 节把有限集的基数定义成集合中的元素个数。定义 4 把基数概念推广到所有集合，包括有限集和无限集。

定义 4 集合 A 和集合 B 有相同的*基数*，当且仅当存在从 A 到 B 的一一对应。

为了看出这个定义与过去的定义（有限集的基数是集合中的元素个数）相一致，注意在任意两个 n 元素有限集之间都存在一一对应，其中 n 是非负整数。

现在把无限集分为两组，一组与自然数集合基数相同，另一组基数不同。

定义 5 有限集或与自然数集基数相同的集合都称为*可数的*。不是可数的集合称为*不可数的*。

现在给出可数集合与不可数集合的例子。

例 18 证明：正奇数集合是可数集合。

解 要证明正奇数集合是可数的，就说明这个集合与正整数集合之间的一一对应。考虑从 $\mathbf{Z}^+$ 到正奇数集合的函数

$$f(n) = 2n-1$$

证明 f 既是单的又是满的，即证明 f 是一一对应的。为了看出 f 是单的，假定 $f(n)=f(m)$。于是 $2n-1=2m-1$，所以 $n=m$。为了看出 f 是满的，假定 t 是正奇数。于是 t 比偶数 $2k$ 少 1，其中 k 是自然数。因此 $t=2k-1=f(k)$。图 3-2 显示这个一一对应。 ■

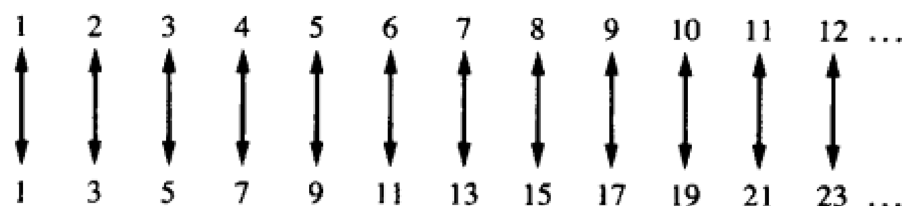


图 3-2 在 $\mathbf{Z}^+$ 和正奇数集合之间的一一对应

无限集是可数的，当且仅当可以把集合中的元素列成序列(下标是正整数)。这是因为从正整数集合到集合 S 的一一对应 f 可表示成序列 a_1，a_2，…，a_n，…，其中 $a_1=f(1)$，$a_2=f(2)$，…，$a_n=f(n)$，…。例如，奇数集合列成序列 a_1，a_2，…，a_n，…，其中 $a_n=2n-1$。

例 19 证明：正有理数集合是可数的。

解 正有理数集合是可数的，这似乎令人惊讶，但下面说明如何把正有理数列成序列 r_1，r_2，…，r_n，…。首先，注意每个正有理数都是两个正整数之商 p/q。这样排列正有理数：在第 1 行列出分母 $q=1$ 的有理数，在第 2 行列出分母 $q=2$ 的有理数，等等，如图 3-3 所示。

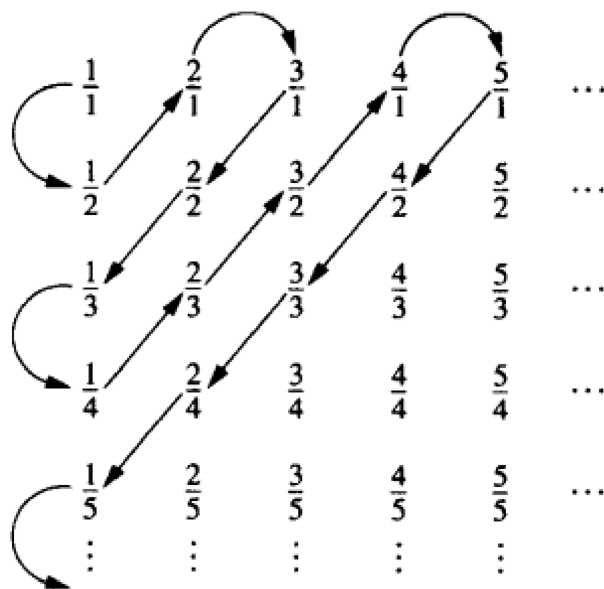


图 3-3 正有理数是可数的

把有理数列成序列的关键是：沿着图 3-3 所示路线，先列出满足 $p+q=2$ 的正有理数 p/q，再列出满足 $p+q=3$ 的正有理数，然后列出满足 $p+q=4$ 的正有理数，等等。每当遇到已经列出过的数 p/q，就不再次列出了。例如，当遇到 $2/2=1$ 时就不列出了，因为已经列出过 $1/1=1$。所构造的正有理数表的初始项是 1，1/2，2，3，1/3，1/4，2/3，3/2，4，5，等等。由于所有有理数都只列出一次(读者可以验证它)，所以已经证明了有理数集合是可数的。 ■

例 20 证明实数集合是不可数的。乔治·康托尔在 1879 年发现了这个事实。下面使用所谓康托尔对角化论证的重要证明方法来证明实数集合是不可数的。在数理逻辑和计算理论中大量地使用这个证明方法。

例 20 证明：实数集合是不可数集合。

解 要证明实数集合是不可数的，就假定实数集合是可数的并得出矛盾。于是，所有落在 0 和 1 之间的实数所成的子集合也是可数的(因为可数集合的所有子集合都是可数的，参见本节末练习 34)。在此假设下，在 0 和 1 之间的实数按照某种顺序列出，比如说 r_1，r_2，r_3，…。设这些实数的十进制表示为

$$
\begin{aligned}
r_1 &= 0.d_{11}d_{12}d_{13}d_{14}\cdots \\
r_2 &= 0.d_{21}d_{22}d_{23}d_{24}\cdots \\
r_3 &= 0.d_{31}d_{32}d_{33}d_{34}\cdots \\
r_4 &= 0.d_{41}d_{42}d_{43}d_{44}\cdots \\
&\vdots
\end{aligned}
$$

其中 $d_{ij}\in\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$。(例如，如果 $r_1=0.237\,941\,02\cdots$，就有 $d_{11}=2$，$d_{12}=3$，$d_{13}=7$，等等。)于是，构造新的实数具有十进制展开式 $r=0.d_1d_2d_3d_4\cdots$，其中用下列规则确定十进制数字：

$$d_i = \begin{cases} 4 & 若\ d_{ii} = 4 \\ 5 & 若\ d_{ii} \neq 4 \end{cases}$$

(作为例子，假定 $r_1=0.237\,941\,02\cdots$，$r_2=0.445\,901\,38\cdots$，$r_3=0.091\,187\,64\cdots$，$r_4=0.805\,539\,00\cdots$，等等，于是 $r=0.d_1d_2d_3d_4\cdots=0.4544\cdots$，其中因为 $d_{11}\neq 4$，$d_1=4$；因为 $d_{22}=4$，$d_2=5$；因为 $d_{33}\neq 4$，$d_3=4$；因为 $d_{44}\neq 4$，$d_4=4$；等等。)

每个实数都有唯一的十进制展开式(展开式结尾全部由数字 9 组成的可能性除外)。于是，实数 r 不等于 r_1，r_2，…中的任何一个，因为对每个 i 来说，r 的十进制展开式与 r_i 的十进制展开式在小数点右边第 i 位是不同的。

由于存在着不在列表中而在 0 和 1 之间的实数 r，所以可以列出在 0 和 1 之间所有实数的假设就必定为假。因此，不能列出在 0 和 1 之间的所有实数，在 0 和 1 之间的实数集合是不可数的。任何含有不可数子集合的集合都是不可数的(参见本节末练习 35)。因此实数集合是不可数的。 ■

练习

1. 求序列$\{a_n\}$的下列各项，其中 $a_n=2\cdot(-3)^n+5^n$。
   a) a_0　　b) a_1　　c) a_4　　d) a_5
2. 如果 a_n 等于下列各值，则序列$\{a_n\}$的项 a_8 是什么？
   a) 2^{n-1}　　b) 7　　c) $1+(-1)^n$　　d) $-(-2)^n$
3. 序列$\{a_n\}$的项 a_0，a_1，a_2 和 a_3 是什么？其中 a_n 等于
   a) 2^n+1　　b) $(n+1)^{n+1}$　　c) $\lfloor n/2\rfloor$　　d) $\lfloor n/2\rfloor+\lceil n/2\rceil$
4. 序列$\{a_n\}$的项 a_0，a_1，a_2 和 a_3 是什么？其中 a_n 等于
   a) $(-2)^n$　　b) 3　　c) $7+4^n$　　d) $2^n+(-2)^n$
5. 列出下列各序列的前 10 项。
   a) 从 2 开始，每个后续项都比前一项多 3 的序列。
   b) 按升序把每个正整数列出 3 次的序列。
   c) 按升序把每个正奇数列出 2 次的序列。
   d) 第 n 项是 $n!-2^n$ 的序列。
   e) 从 3 开始，每个后续项都是前一项的 2 倍的序列。
   f) 前两项都是 1，每个后续项都是前两项之和的序列。(这是著名的斐波那契序列，后文有述。)
   g) 第 n 项是数 n 的二进制展开式(在 2.5 节定义)的位数的序列。
   h) 第 n 项是第 n 个英文单词中包含的字母数的序列。
6. 列出下列各序列的前 10 项。
   a) 从 10 开始，每个后项都是从前项减去 3 所得的序列。
   b) 第 n 项是前 n 个正整数之和的序列。
   c) 第 n 项是 3^n-2^n 的序列。
   d) 第 n 项是$\lfloor\sqrt{n}\rfloor$的序列。
   e) 前两项是 1 和 2，每个后续项都是前两项之和的序列。
   f) 第 n 项是二进制展开式(在 2.5 节定义)有 n 位的最大整数的序列(用十进制记号写答案)。
   g) 如下顺序地构造各项的序列：从 1 开始，然后加 1，然后乘 1，然后加 2，然后乘 2，等等。
   h) 第 n 项是使得 $k!\leqslant n$ 的最大整数 k 的序列。
7. 求至少 3 个不同的序列，初始项都是 1，2，4，用简单的公式或规则产生各项。
8. 求至少 3 个不同的序列，初始项都是 3，5，7，用简单的公式或规则产生各项。
9. 对于下列每个整数表，给出简单的公式或规则，以产生从给定的表开始的整数序列项。
   a) 1，0，1，1，0，0，1，1，1，0，0，0，1，…
   b) 1，2，2，3，4，4，5，6，6，7，8，8，…

c) 1, 0, 2, 0, 4, 0, 8, 0, 16, 0, …
d) 3, 6, 12, 24, 48, 96, 192, …
e) 15, 8, 1, −6, −13, −20, −27, …
f) 3, 5, 8, 12, 17, 23, 30, 38, 47, …
g) 2, 16, 54, 128, 250, 432, 686, …
h) 2, 3, 7, 25, 121, 721, 5041, 40 321, …

10. 对于下列每个整数表，给出简单的公式或规则，以产生从给定的表开始的整数序列项。

a) 3, 6, 11, 18, 27, 38, 51, 66, 83, 102, …
b) 7, 11, 15, 19, 23, 27, 31, 35, 39, 43, …
c) 1, 10, 11, 100, 101, 110, 111, 1000, 1001, 1010, 1011, …
d) 1, 2, 2, 2, 3, 3, 3, 3, 3, 5, 5, 5, 5, 5, 5, 5, …
e) 0, 2, 8, 26, 80, 242, 728, 2186, 6560, 19 682, …
f) 1, 3, 15, 105, 945, 10 395, 135 135, 2 027 025, 34 459 425, …
g) 1, 0, 0, 1, 1, 1, 0, 0, 0, 0, 1, 1, 1, 1, 1, …
h) 2, 4, 16, 256, 65 536, 4 294 967 296, …

*11. 证明：如果 a_n 表示不是完全平方的第 n 个正整数，则 $a_n=n+\{\sqrt{n}\}$，其中 $\{x\}$ 表示最接近于实数 x 的整数。

*12. 设 a_n 表示序列 1, 2, 2, 3, 3, 3, 4, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, …的第 n 项，构造这个序列的方法是包含整数 k 恰好 k 次。证明：$a_n=\lfloor\sqrt{2n}+1/2\rfloor$。

13. 下列各和的值是什么？

a) $\sum_{k=1}^{5}(k+1)$ b) $\sum_{j=0}^{4}(-2)^j$ c) $\sum_{i=1}^{10}3$ d) $\sum_{j=0}^{8}(2^{j+1}-2^j)$

14. 下列各和的值是什么?其中 $S=\{1,3,5,7\}$。

a) $\sum_{j\in S}j$ b) $\sum_{j\in S}j^2$ c) $\sum_{j\in S}(1/j)$ d) $\sum_{j\in S}1$

15. 下列每个几何序列这些项之和的值是什么？

a) $\sum_{j=0}^{8}3\cdot 2^j$ b) $\sum_{j=1}^{8}2^j$ c) $\sum_{j=2}^{8}(-3)^j$ d) $\sum_{j=0}^{8}2\cdot(-3)^j$

16. 求下列各和之值。

a) $\sum_{j=0}^{8}(1+(-1)^j)$ b) $\sum_{j=0}^{8}(3^j-2^j)$ c) $\sum_{j=0}^{8}(2\cdot 3^j+3\cdot 2^j)$ d) $\sum_{j=0}^{8}(2^{j+1}-2^j)$

17. 计算下列各双重和

a) $\sum_{i=1}^{2}\sum_{j=1}^{3}(i+j)$ b) $\sum_{i=0}^{2}\sum_{j=0}^{3}(2i+3j)$ c) $\sum_{i=1}^{3}\sum_{j=0}^{2}i$ d) $\sum_{i=0}^{2}\sum_{j=1}^{3}ij$

18. 计算下列各双重和

a) $\sum_{i=1}^{3}\sum_{j=1}^{2}(i-j)$ b) $\sum_{i=0}^{3}\sum_{j=0}^{2}(3i+2j)$ c) $\sum_{i=1}^{3}\sum_{j=0}^{2}j$ d) $\sum_{i=0}^{2}\sum_{j=0}^{3}i^2j^3$

19. 证明：$\sum_{j=1}^{n}(a_j-a_{j-1})=a_n-a_0$，其中 $a_0,a_1,\cdots,a_n$ 是实数序列。这种类型的和称为迭进。

20. 利用恒等式 $1/(k(k+1))=1/k-1/(k+1)$ 和练习 19 来计算 $\sum_{k=1}^{n}1/(k(k+1))$。

21. 从 $k=1$ 到 $k=n$ 对恒等式 $k^2-(k-1)^2=2k-1$ 两边求和，并且利用练习 19，求：

a) $\sum_{k=1}^{n}(2k-1)$ 的公式(前 n 个奇自然数之和)。 b) $\sum_{k=1}^{n}k$ 的公式。

*22. 利用练习 19 给出的技术以及练习 21 中(b) 的结果，求 $\sum_{k=1}^{n}k^2$ 的公式。

23. 用表 3-2 求 $\sum_{k=100}^{200}k$。

24. 用表 3-2 求 $\sum_{k=99}^{200} k^3$。

*25. 当 m 是正整数时，求 $\sum_{k=0}^{m} \lfloor \sqrt{k} \rfloor$ 的公式。（提示：用 $\sum_{k=1}^{n} k^2$ 的公式。）

*26. 当 m 是正整数时，求 $\sum_{k=0}^{m} \lfloor \sqrt[3]{k} \rfloor$ 的公式。（提示：用 $\sum_{k=1}^{n} k^3$ 的公式。）

对于乘积也有一个特殊记号。$a_m, a_{m+1}, \cdots, a_n$ 的乘积表示为 $\prod_{j=m}^{n} a_j$。

27. 下列乘积的值是什么？

a) $\prod_{i=0}^{10} i$　　b) $\prod_{i=5}^{8} i$　　c) $\prod_{i=1}^{100} (-1)^i$　　d) $\prod_{i=1}^{10} 2$

回忆一下，阶乘函数在正整数 n 上的值（表示成 $n!$）是从 1 到 n 的正整数之积。另外规定 $0! = 1$。

28. 用乘积记号表示 $n!$。

29. 求 $\sum_{j=0}^{4} j!$。

30. 求 $\prod_{j=0}^{4} j!$。

31. 确定下列各集合是否可数或不可数。对可数的集合，说明在自然数集合和该集合之间的一一对应。

a)负整数　　b)偶数　　c)0 和 1/2 之间的实数　　d)7 的整倍数

*32. 确定下列各集合是否可数或不可数。对可数的集合，说明在自然数集合和该集合之间的一一对应。

a)不能被 3 整除的整数　　b)能被 5 整除但不能被 7 整除的整数

c)十进制表示中只含数字 1 的实数　　d)十进制表示中只含数字 1 或 9 的实数

33. 如果 A 是不可数集合而 B 是可数集合，那么 $A-B$ 一定不可数吗？

34. 证明：可数集合的子集合也是可数的。

35. 证明：如果 A 是不可数集合并且 $A \subseteq B$，则 B 是不可数的。

36. 证明：两个可数集合的并集是可数的。

**37. 证明：可数的多个可数集合的并集是可数的。

38. 证明：集合 $Z^+ \times Z^+$ 是可数的。

*39. 证明：所有位串的集合是可数的。

*40. 证明：二次方程 $ax^2+bx+c=0$ 的实数解集合是可数的，其中 a、b 和 c 都是整数。

*41. 证明：用具体的程序设计语言编写的所有计算机程序的集合是可数的。（提示：可以认为用程序设计语言编写的计算机程序是有穷字母表上的符号串。）

*42. 证明：从正整数到集合{0，1，2，3，4，5，6，7，8，9}的函数的集合是不可数的。（提示：首先建立在 0 到 1 之间实数集与这些函数的子集之间的一一对应。为此让实数 $0.d_1d_2\cdots d_n\cdots$ 对应到函数 f，满足 $f(n)=d_n$。）

*43. 如果存在着求函数值的计算机程序，就说这个函数是可计算的。用练习 41 和练习 42 证明：存在着不可计算的函数。

*44. 证明：正有理数集是可数的。方法如下。建立一个函数把满足 $\gcd(p, q)=1$ 的有理数 p/q 如下对应到构成的 11 进制数：先是 p 的十进制表示，然后接着 11 进制数字 A（A 对应于十进制数 10），然后接着 q 的十进制表示。

*45. 证明：正有理数集是可数的。方法如下。假设 $K(m/n) = p_1^{2a_1} p_2^{2a_2} \cdots p_s^{2a_s} q_1^{2b_1-1} q_2^{2b_2-1} \cdots q_t^{2b_t-1}$，其中 $\gcd(m, n)=1$并且 m 和 n 的素数幂因子分解是 $m = p_1^{a_1} p_2^{a_2} \cdots p_s^{a_s}$ 和$n = q_1^{b_1} q_2^{b_2} \cdots q_t^{b_t}$，证明函数 K 是正有理数集合和正整数集合之间的一一对应。

3.3 数学归纳法

3.3.1 引言

前 n 个正奇数之和的公式是什么？对 $n=1$，2，3，4，5 来说前 n 个正奇数之和是

$$1=1,\quad 1+3=4,\quad 1+3+5=9,$$
$$1+3+5+7=16,\quad 1+3+5+7+9=25。$$

根据这些值，有理由猜测前 n 个正奇数之和是 n^2。需要一个方法证明这个猜测是正确的，如果事实上猜测正确的话。

数学归纳法是证明这种断言的极端重要的证明技术。在本节和后几章中，数学归纳法将被大量用来证明关于各种各样离散对象的结果。例如用来证明关于算法的复杂性、特定类型计算机程序的正确性、有关图与树的定理，以及各种恒等式和不等式的结果。

本节描述如何使用数学归纳法和为什么数学归纳法是有效的证明技术。一定要注意的是：数学归纳法只能证明通过其他方式获得的结果，它不是发现公式或定理的工具。

对数学归纳法有几种有用的解释，能帮助记忆这个原理是如何工作的。一种解释是有一队人，甲、乙、丙，等等。每个人听到秘密后，都把它告诉队中下一个人。设 $P(n)$ 是命题：第 n 个人知道秘密。于是 $P(1)$ 为真，因为把秘密告诉了甲；$P(2)$ 为真，因为甲把秘密告诉了乙；$P(3)$ 为真，因为乙把秘密告诉了丙；等等。根据数学归纳法原理，队中每个人都知道了秘密。图 3-4 解释了这一点。（当然，假定了每个人都原封不动地把秘密传给下一个人，这通常与实际生活不符。）

解释数学归纳法的另一种方式是考虑无穷长的一行多米诺骨牌，标记成 1，2，3，…，n，其中每张多米诺骨牌都直立着。设 $P(n)$ 是命题：多米诺骨牌 n 被撞倒。如果第一张多米诺骨牌被撞倒，即 $P(1)$ 为真，并且如果每当第 n 张多米诺骨牌被撞倒时，它也撞倒第 $n+1$ 张多米诺骨牌，即如果 $P(n)\to P(n+1)$ 为真，那么所有的多米诺骨牌都被撞倒。图 3-5 解释这一点。



图 3-4 传递秘密的人们

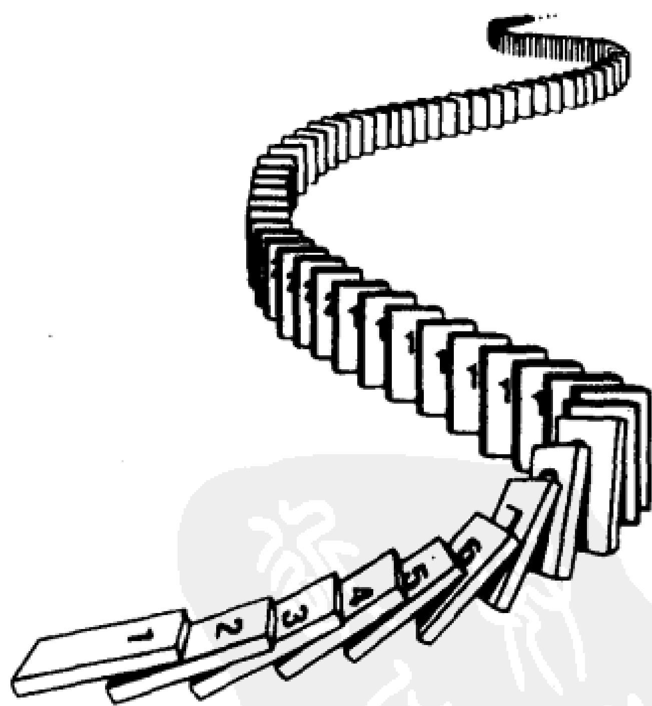


图 3-5 用多米诺骨牌解释数学归纳法如何工作

3.3.2 数学归纳法

许多定理说：对所有正整数 n 来说，$P(n)$ 为真，其中 $P(n)$ 是命题函数，比方说是命题 $1+2+\cdots+n=n(n+1)/2$ 或命题 $n\leqslant 2^n$。数学归纳法是证明这种类型的定理的技术。换句话说，

数学归纳法用来证明形如$\forall nP(n)$的命题，其中论域是正整数集。

通过数学归纳法证明对每个正整数 n 来说，$P(n)$为真，一共包含两个步骤：

基础步骤：证明命题 $P(1)$为真。

归纳步骤：证明对每个正整数 k 来说，蕴含式 $P(k)\to P(k+1)$为真。

这里对固定的正整数 k 来说，命题 $P(k)$称为*归纳假设*。当完成了数学归纳法证明的两个步骤时，就证明了对所有正整数 n 来说，$P(n)$为真；即证明了$\forall nP(n)$为真。

把这种证明技术表示成推理规则就是

$$[P(1)\land\forall k(P(k)\to P(k+1))]\to\forall nP(n)$$

由于数学归纳法是如此重要的证明技术，所以值得详细解释使用这个技术的证明步骤。为了证明对所有正整数 n 来说，$P(n)$为真，首先证明 $P(1)$为真。这等于证明当在 $P(n)$里用 1 替换 n 时所得到的特殊命题为真。然后必须证明对每个正整数 k 来说，都有 $P(k)\to P(k+1)$为真。为了证明对每个正整数 k 来说这个蕴含式为真，需要证明当 $P(k)$为真时 $P(k+1)$不能为假。可以通过假设 $P(k)$为真，而且证明在此假设下 $P(k+1)$也必然为真，来完成这个证明。

注意 *在数学归纳法证明里并不假定对所有正整数来说 $P(k)$为真！只是证明了：若假定 $P(k)$为真，则 $P(k+1)$也为真。因此，数学归纳法证明不属于回避问题或循环论证的情形。*

当用数学归纳法来证明定理时，首先证明 $P(1)$为真。然后知道 $P(2)$为真，因为$P(1)$蕴含$P(2)$。另外，还知道 $P(3)$为真，因为 $P(2)$蕴含 $P(3)$。以这样的方式继续下去，就可以看出对任意正整数 n 来说，$P(n)$为真。⊖

3.3.3 数学归纳法证明的例子

下面要用各种例子来说明如何用数学归纳法证明定理。首先证明前 n 个正奇数之和的公式。（在本节里用数学归纳法证明的许多定理都可以用不同的方法来证明。不过，用多种方式来证明一个定理是值得尝试的，因为一种证明方法可能成功而另外一种方法或许不成功。）

例 1 用数学归纳法证明：前 n 个正奇数之和是 n^2。

解 设 $P(n)$表示命题：前 n 个正奇数之和是 n^2。必须首先完成基础步骤；即必须证明 $P(1)$为真。然后必须完成归纳步骤；即必须证明当假定 $P(k)$为真时 $P(k+1)$为真。

基础步骤：根据 $P(1)$有：前 1 个正奇数之和是 1^2。这是真的，因为第 1 个正奇数是 1。

归纳步骤：为了完成归纳步骤，必须证明对每个正整数 n 来说，命题 $P(k)\to P(k+1)$为真。为了做到这一点，假定对正整数 k 来说 $P(k)$为真；即

$$1+3+5+\cdots+(2k-1)=k^2$$

（注意第 k 个正奇数是$(2k-1)$，因为这个整数是通过向 1 加 $k-1$ 次 2 所获得的。）必须证明假定 $P(k)$为真，则 $P(k+1)$为真。注意 $P(k+1)$是命题

$$1+3+5+\cdots+(2k-1)+(2k+1)=(k+1)^2$$

⊖ 历史注记：已知最早的对数学归纳法的使用，是在 16 世纪数学家弗朗西斯科·毛洛利可（Francesco Maurolico，1494—1575）的著作里。毛洛利可写过大量的关于经典数学的著作，并且对几何学和光学做出过许多贡献。在他的著作《*Arithmeticorum Libri Duo*》里，毛洛利可给出了整数的各种性质和对这些性质的证明。为了证明其中的某些性质，他设计出数学归纳法这个方法。在这本书里他对数学归纳法的第一次使用是为了证明前 n 个正奇数之和等于 n^2。

所以，假定 $P(k)$ 为真，得出

$$\begin{aligned}1+3+5+\cdots+(2k-1)+(2k+1) &= [1+3+5+\cdots+(2k-1)]+(2k+1)\\ &= k^2+(2k+1)\\ &= k^2+2k+1\\ &= (k+1)^2\end{aligned}$$

这样证明了从 $P(k)$ 得出 $P(k+1)$。注意在第二个等式里使用了归纳假设 $P(k)$，以便用 k^2 来代替前 k 个正奇数之和。

因为 $P(1)$ 为真，而且对所有正整数 k 来说，蕴含式 $P(k)\to P(k+1)$ 为真，所以数学归纳法原理就证明了对所有正整数 n 来说 $P(n)$ 为真。 ■

例 2 用数学归纳法来证明一个不等式。

例 2 用数学归纳法证明：对所有正整数 n 来说，有不等式

$$n<2^n$$

解 设 $P(n)$ 是命题"$n<2^n$"。

基础步骤：$P(1)$ 为真，因为 $1<2^1=2$。

归纳步骤：假定对正整数 k 来说，$P(k)$ 为真；即假定 $k<2^k$。需要证明 $P(k+1)$ 为真。即需要证明 $k+1<2^{k+1}$。向 $k<2^k$ 的两边都加 1，然后注意 $1\leqslant 2^k$，这样就给出

$$k+1<2^k+1\leqslant 2^k+2^k=2^{k+1}$$

在 $P(k)$ 为真的假定的基础上，已经证明了 $P(k+1)$ 为真，即 $k+1<2^{k+1}$。归纳步骤完毕。

因此，根据数学归纳法原理，已经证明对所有正整数 n 来说，$n<2^n$ 为真。 ■

下面，用数学归纳法证明一个与整数整除性有关的定理。

例 3 用数学归纳法证明：每当 n 是正整数时，n^3-n 就被 3 整除。

解 为了构造这个证明，设 $P(n)$ 是命题"n^3-n 被 3 整除"。

基础步骤：$P(1)$ 为真，因为 $1^3-1=0$ 被 3 整除。

归纳步骤：假定 $P(k)$ 为真；即 k^3-k 被 3 整除。必须证明 $P(k+1)$ 为真。即必须证明 $(k+1)^3-(k+1)$ 被 3 整除。注意

$$\begin{aligned}(k+1)^3-(k+1) &= (k^3+3k^2+3k+1)-(k+1)\\ &= (k^3-k)+3(k^2+k)\end{aligned}$$

因为在这个和里的两项都被 3 整除(第一项是根据归纳步骤的假设，第二项是因为它是一个整数的 3 倍)，由此得出 $(k+1)^3-(k+1)$ 也被 3 整除。这样就完成了归纳步骤。因此，根据数学归纳法原理，每当 n 是正整数时，n^3-n 就被 3 整除。 ■

有时需要证明对 $n=b$，$b+1$，$b+2$，…来说 $P(n)$ 为真，其中 b 是不等于 1 的整数。只要改变基础步骤，就可以用数学归纳法来完成这个证明。例如，考虑例 4，它证明对所有非负整数来说，一个求和公式是有效的，所以需要证明对 $n=0$，1，2，…，$P(n)$ 为真。

例 4 用数学归纳法证明：对所有非负整数来说

$$1+2+2^2+\cdots+2^n=2^{n+1}-1$$

解 设 $P(n)$ 是命题：对整数 n 来说，这个公式正确。

基础步骤：$P(0)$ 为真，因为 $2^0=1=2^1-1$。

归纳步骤：假定 $P(k)$ 为真。为了利用这个假定完成归纳步骤，必须证明 $P(k+1)$ 为真，即

$$1+2+2^2+\cdots+2^k+2^{k+1}=2^{(k+1)+1}-1=2^{k+2}-1$$

使用归纳假设 $P(k)$，得出

$$1+2+2^2+\cdots+2^k+2^{k+1}=(1+2+2^2+\cdots+2^k)+2^{k+1}$$

$$= (2^{k+1} - 1) + 2^{k+1} = 2 \cdot 2^{k+1} - 1$$
$$= 2^{k+2} - 1$$

这样就完成了归纳步骤，证毕。■

例4说明，为了用数学归纳法证明对 $n=b$，$b+1$，$b+2$，…来说 $P(n)$ 为真，其中 b 是不等于1的整数，就证明 $P(b)$ 为真(基础步骤)，然后证明对 $k=b$，$b+1$，$b+2$，…来说，蕴含式 $P(k) \to P(k+1)$ 为真(归纳步骤)。注意 b 可以是非负的、零或正的。根据前面用过的多米诺骨牌的比喻，想像首先撞倒第 b 张多米诺骨牌(基础步骤)，当每张多米诺骨牌倒下时，它就撞倒下一张多米诺骨牌(归纳步骤)。请读者去证明这种形式的归纳是有效的(见练习76)。

例4里给出的公式是几何序列各项之和的一般结果(3.2节定理1)的特殊情形。我们将用数学归纳法给出这个公式的另一种证明。

例5 几何序列求和 用数学归纳法证明几何序列有穷多项之和的下列公式：

$$\sum_{j=0}^{n} ar^j = a + ar + ar^2 + \cdots + ar^n = \frac{ar^{n+1} - a}{r - 1}, \text{当 } r \neq 1 \text{ 时}$$

解 为了用数学归纳法来证明这个公式，设 $P(n)$ 是命题：这个公式里的几何序列前 $n+1$ 项之和是正确的。

基础步骤：$P(0)$ 为真，因为 $a = \frac{ar - a}{r - 1}$。

归纳步骤：假定 $P(k)$ 为真。即假定

$$a + ar + ar^2 + \cdots + ar^k = \frac{ar^{k+1} - a}{r - 1}$$

为了证明它蕴含 $P(k+1)$ 为真，向这个等式的两边都加上 ar^{k+1} 得到

$$a + ar + ar^2 + \cdots + ar^k + ar^{k+1} = \frac{ar^{k+1} - a}{r - 1} + ar^{k+1}$$

改写这个等式的右边可得

$$\frac{ar^{k+1} - a}{r - 1} + ar^{k+1} = \frac{ar^{k+1} - a}{r - 1} + \frac{ar^{k+2} - ar^{k+1}}{r - 1} = \frac{ar^{k+2} - a}{r - 1}$$

把这些等式组合起来就给出

$$a + ar + ar^2 + \cdots + ar^k + ar^{k+1} = \frac{ar^{k+2} - a}{r - 1}$$

这样就证明了若 $P(k)$ 为真则 $P(k+1)$ 必然也为真。这样就完成了归纳论证并且证明了几何序列各项之和的公式是正确的。■

前面指出过，例4里的公式是例5里的公式当 $a=1$ 和 $r=2$ 时的情形。读者应当验证在一般公式里设 a 和 r 取这些值，就给出与在例4里同样的公式。

在下一个例子里将证明一组正整数的倒数之和的一个重要不等式。

例6 调和数不等式 对 $j=1$，2，3，…，调和数 H_j 定义为

$$H_j = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{j}$$

例如，

$$H_4 = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = \frac{25}{12}$$

用数学归纳法证明：每当 n 是非负整数时就有

$$H_{2^n} \geqslant 1 + \frac{n}{2}$$

解 为了完成这个证明，设 $P(n)$ 是命题：$H_{2^n} \geqslant 1+n/2$。

基础步骤：$P(0)$ 为真，因为 $H_{2^0} = H_1 = 1 \geqslant 1+0/2$。

归纳步骤：假定 $P(k)$ 为真，所以 $H_{2^k} \geqslant 1+k/2$。必须证明在这个假定之下 $P(k+1)$ 也必然为真，根据 $P(k+1)$ 有：$H_{2^{k+1}} \geqslant 1+(k+1)/2$。这是可以证明的，因为

$$\begin{aligned}
H_{2^{k+1}} &= 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{2^k} + \frac{1}{2^k+1} + \cdots + \frac{1}{2^{k+1}} && \text{（根据调和数的定义）} \\
&= H_{2^k} + \frac{1}{2^k+1} + \cdots + \frac{1}{2^{k+1}} && \text{（根据调和数的定义）} \\
&\geqslant \left(1+\frac{k}{2}\right) + \frac{1}{2^k+1} + \cdots + \frac{1}{2^{k+1}} && \text{（根据归纳假设）} \\
&\geqslant \left(1+\frac{k}{2}\right) + 2^k \cdot \frac{1}{2^{k+1}} && \text{（因为有 } 2^k \text{ 项都不小于 } 1/2^{k+1}\text{）} \\
&\geqslant \left(1+\frac{k}{2}\right) + \frac{1}{2} \\
&= 1 + \frac{k+1}{2}
\end{aligned}$$

这样就完成了证明的归纳步骤。因此，对所有非负整数 n 来说，调和数不等式都是有效的。 ■

注意 *可以用这里证明的不等式去证明调和级数*

$$1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots$$

是发散的无穷级数。在无穷级数的研究里这是一个重要的例子。

下一个例子说明如何用数学归纳法去验证有穷集合的子集个数的公式。

例 7 有穷集合的子集个数 用数学归纳法证明：若 S 是有 n 个元素的有穷集合，则 S 有 2^n 个子集。（在第 4 章里我们将以多种方式直接地证明这个结果。）

解 设 $P(n)$ 是命题：有 n 个元素的集合有 2^n 个子集。

基础步骤：$P(0)$ 为真，因为有 0 个元素的集合，即空集，恰有 $2^0=1$ 个子集，因为它有一个子集，即它自身。

归纳步骤：假定 $P(k)$ 为真，即每个有 k 个元素的集合都有 2^k 个子集。必须证明在这个假定下命题 $P(k+1)$ 也必然为真，即：每个有 $k+1$ 个元素的集合都有 2^{k+1} 个子集。为了证明它，设 T 是有 $k+1$ 个元素的集合。则可以写出 $T=S\cup\{a\}$，其中 a 是 T 中的一个元素而 $S=T-\{a\}$。可以用下面的方式来获得 T 的子集。对 S 的每个子集 X 来说，恰存在两个 T 的子集，即 X 和 $X\cup\{a\}$（在图 3-6 对此进行解释）。这些集合构成了 T 的所有子集并且都互不相等。因为 S 有 2^k 个子集，所以 T 有 $2\cdot 2^k=2^{k+1}$ 个子集。这样就完成了归纳论证。 ■

例 8 证明：若 n 是正整数，则 $1+2+\cdots+n=n(n+1)/2$。

解 设 $P(n)$ 是命题：前 n 个正整数之和是 $n(n+1)/2$。必须做两件事情来证明对 $n=1, 2, 3, \cdots$，$P(n)$ 为真。即必须证明 $P(1)$ 为真，以及对 $k=1, 2, 3, \cdots$ 来说蕴含式 $P(k)$ 蕴含 $P(k+1)$ 为真。

基础步骤：$P(0)$ 为真，因为 $1=1(1+1)/2$。

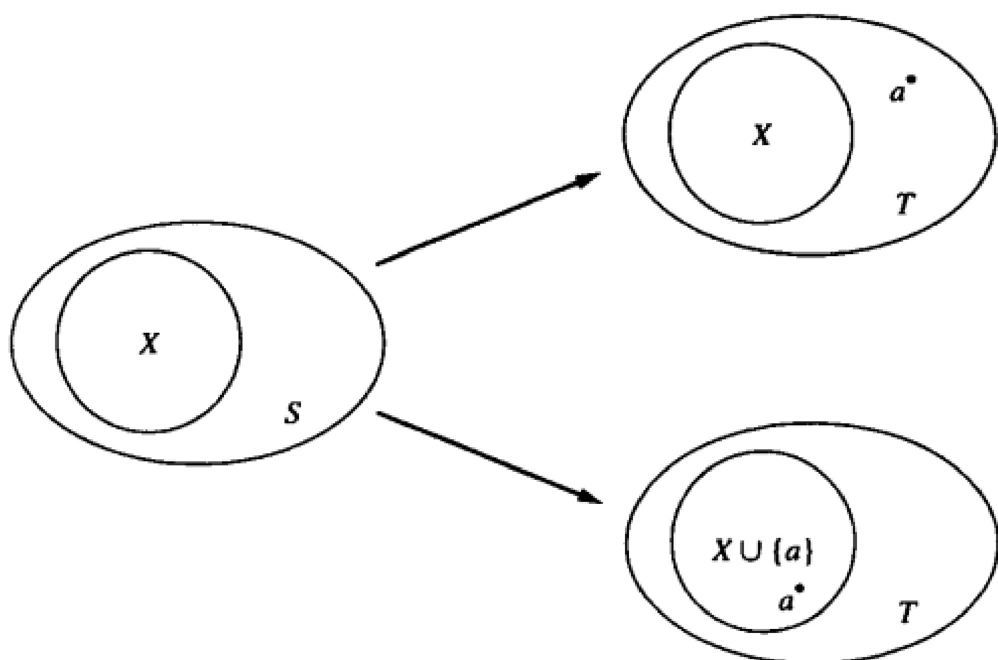


图 3-6 生成有 $k+1$ 个元素的集合的子集。这里 $T=S\cup\{a\}$

归纳步骤：假定 $P(k)$成立，使得

$$1+2+\cdots+k=k(k+1)/2$$

在这个假定之下，必须证明 $P(k+1)$为真，即

$$1+2+\cdots+k+(k+1)=(k+1)[(k+1)+1]/2=(k+1)(k+2)/2$$

也为真。向 $P(k)$里的等式的两边都加上 $k+1$ 就得到

$$\begin{aligned}1+2+\cdots+k+(k+1)&=k(k+1)/2+(k+1)\\&=[(k/2)+1](k+1)\\&=(k+1)(k+2)/2\end{aligned}$$

最后这个等式证明 $P(k+1)$为真。这样就完成了归纳步骤，证毕。■

例 9 用数学归纳法证明：对每个满足 $n\geqslant 4$ 的正整数 n 来说，有 $2^n<n!$。

解 设 $P(n)$是命题：$2^n<n!$。

基础步骤：为了证明对 $n\geqslant 4$ 来说这个不等式成立，就需要让基础步骤是 $P(4)$。注意 $P(4)$ 为真，因为 $2^4=16<4!=24$。

归纳步骤：假定 $P(k)$为真，即假定 $2^k<k!$。必须证明 $P(k+1)$为真。即必须证明 $2^{k+1}<(k+1)!$。向不等式 $2^k<k!$ 的两边都乘上 2 就得到

$$2\cdot 2^k<2\cdot k!<(k+1)\cdot k!=(k+1)!$$

这说明当 $P(k)$为真时 $P(k+1)$为真。这样就完成了证明的归纳步骤。因此，就得出对每个满足 $n\geqslant 4$的正整数 n 来说，$2^n<n!$ 为真。■

例 10 用数学归纳法证明下述对德·摩根律之一的推广：每当 A_1，A_2，…，A_n 都是全集 U 的子集并且 $n\geqslant 2$ 时，就有

$$\overline{\bigcap_{j=1}^{n} A_j}=\bigcup_{j=1}^{n}\overline{A_j}$$

解 设 $P(n)$是对 n 个集合来说的上述恒等式。

基础步骤：命题 $P(2)$断言$\overline{A_1\cap A_2}=\overline{A_1}\cup\overline{A_2}$。这是德·摩根律之一；在 1.5 节里证明过它。

归纳步骤：假定 $P(k)$为真，即每当 A_1，A_2，…，A_k 都是全集 U 的子集时，就有

$$\overline{\bigcap_{j=1}^{k} A_j} = \bigcup_{j=1}^{k} \overline{A_j}$$

为了完成归纳步骤，必须证明：若对 U 的任意 k 个子集来说，这个等式成立，则对 U 的任意 $k+1$ 个子集来说，它也必然是有效的。假定 A_1，A_2，…，A_k，A_{k+1} 都是 U 的子集。当假定归纳假设成立时，就得出

$$\begin{aligned}\overline{\bigcap_{j=1}^{k+1} A_j} &= \overline{\left(\bigcap_{j=1}^{k} A_j\right) \cap A_{k+1}} \\ &= \left(\overline{\bigcap_{j=1}^{k} A_j}\right) \cup \overline{A_{k+1}} && \text{（根据德·摩根律）} \\ &= \left(\bigcup_{j=1}^{k} \overline{A_j}\right) \cup \overline{A_{k+1}} && \text{（根据归纳假设）} \\ &= \bigcup_{j=1}^{k+1} \overline{A_j}\end{aligned}$$

这样就完成了归纳法证明。■



例 11 说明如何用数学归纳法证明关于用形如字母“L”的碎片覆盖棋盘的结果。

例 11 设 n 是正整数。证明：可以用 L 形状的碎片来铺满去掉 1 个格的任何 $2^n \times 2^n$ 棋盘，其中这些碎片一次覆盖 3 个格，如图 3-7 所示。

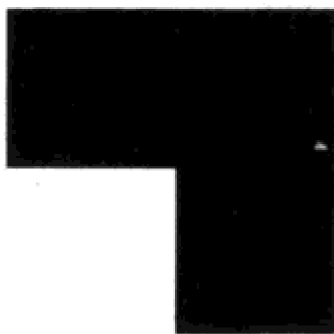


图 3-7 一个 L 形状的碎片

解 设 $P(n)$ 是命题：可以用 L 形状的碎片来铺满去掉 1 个格的任何 $2^n \times 2^n$ 棋盘。可以用数学归纳法证明对所有正整数 n 来说 $P(n)$ 为真。

基础步骤：命题 $P(1)$ 为真，因为可以用 1 个 L 形状的碎片来铺满去掉 1 个格的任何 2×2 棋盘，如图 3-8 所示。

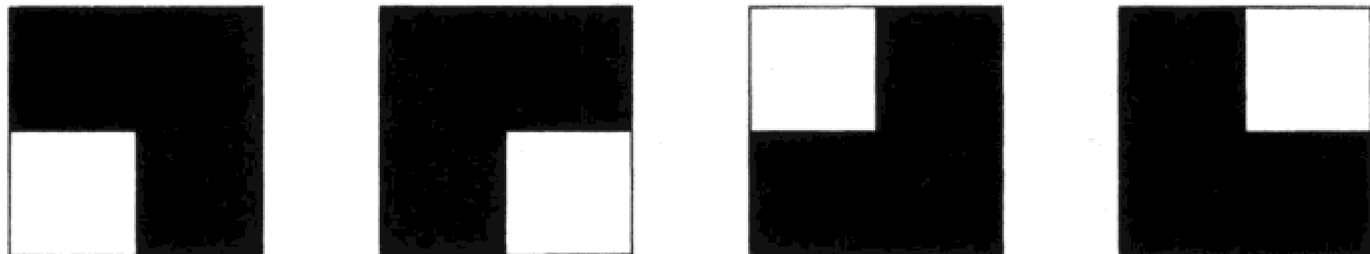


图 3-8 铺满去掉 1 个格的 2×2 棋盘

归纳步骤：假定 $P(k)$ 为真；即假定可以用 L 形状的碎片来铺满去掉 1 个格的任何 $2^k \times 2^k$ 棋盘。必须证明在这个假定下 $P(k+1)$ 也必然为真；即可以用 L 形状的碎片来铺满去掉 1 个格的任何 $2^{k+1} \times 2^{k+1}$ 棋盘。

为了看出这一点，考虑去掉 1 个格的 $2^{k+1} \times 2^{k+1}$ 棋盘。通过在两个方向上都一分为二，把这个棋盘分成 $2^k \times 2^k$ 大小的 4 个棋盘。在图 3-9 里对此进行说明。这 4 个棋盘中有 3 个没有去掉任何格。第 4 个 $2^k \times 2^k$ 棋盘去掉了 1 个格，所以根据归纳假设，可以用 L 形状的碎片来覆盖它。现在暂时去掉另外 3 个 $2^k \times 2^k$ 棋盘的一角上的格，这些角是原来较大的棋盘的中心，如图 3-10 所示。根据归纳假设，可以用 L 形状的碎片来铺满这 3 个去掉了 1 个格的 $2^k \times 2^k$ 棋盘中的每一个。另外，可以用 1 个 L 形状的碎片来铺满这 3 个暂时去掉的格。因此，可以用 L 形状的碎片来铺满整个的 $2^{k+1} \times 2^{k+1}$ 棋盘。证毕。■

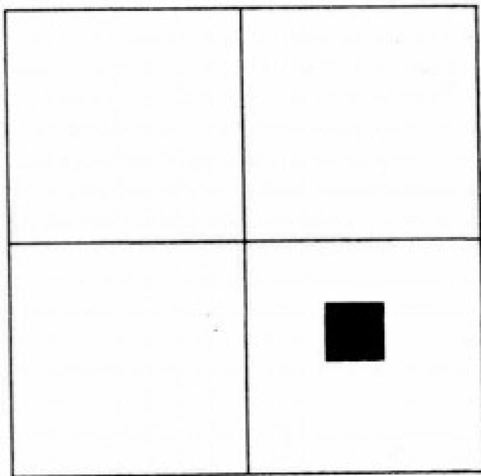


图3-9　把一个 $2^{k+1}\times 2^{k+1}$ 棋盘分成 4 个 $2^k\times 2^k$ 棋盘

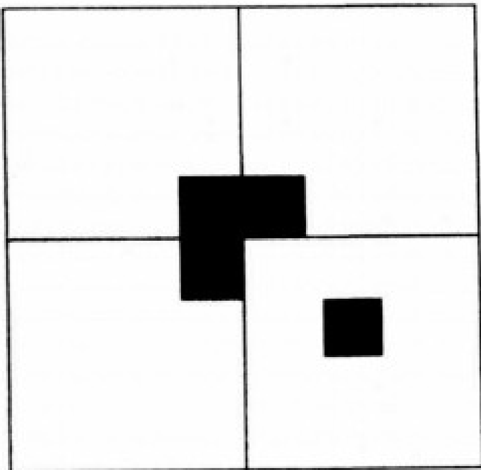


图 3-10　铺满去掉了 1 个格的 $2^{k+1}\times 2^{k+1}$ 棋盘

例 12　可以用贪心算法来调度在单间演讲厅里举行的 m 场报告 t_1，t_2，…，t_m 的子集合。假设报告 t_j 在时间 b_j 开始并在时间 e_j 结束。(两个报告不能同时进行，一个报告可以在另一个报告结束时开始。)假设按照结束时间非降的顺序列出报告，得到 $e_1\leqslant e_2\leqslant\cdots\leqslant e_m$。贪心算法这样进行：在每个阶段，从所有已经安排好的报告结束之后才开始的那些报告中，选择具有最早结束时间的报告(这个算法总是加入具有最早结束时间的报告)。将要证明这个贪心算法在下列意义下是最优的：这个算法总是尽可能安排最多的报告。为了证明这个算法的最优性，对变量 n 使用数学归纳，n 是这个算法安排的报告数。设 $P(n)$是命题：如果贪心算法安排 n 个报告，则不可能安排超过 n 个报告。

基础步骤：假设贪心算法只安排了一个报告 t_1。这意味着每个其他报告不能在 e_1 之后开始。否则，按照结束时间非降的顺序考虑报告，就应该加入所遇到的第一个这样的报告。因此，在时间 e_1，每个剩余的报告都要求使用演讲厅，因为这些报告都在 e_1 或 e_1 之前开始并且在 e_1 之后结束。结论是不能安排两个报告，因为两个报告都在时间 e_1 要求使用演讲厅。这样就证明 $P(1)$为真并完成基础步骤。

归纳步骤：假设 $P(k)$为真，也就是说，给定任意的报告集合(无论有多大)，当贪心算法选择了 k 个报告时，贪心算法总是安排最多可能的报告。当贪心算法安排 k 个报告时，贪心算法总是产生最优解，在这个假设下，必须证明贪心算法已经选择了最大可能数目的报告。也就是说，假设 $P(k)$为真，需要证明 $P(k+1)$为真。

为了完成归纳步骤，首先证明存在一种包含最多可能的报告的安排，这个安排包含一个具有最早结束时间的报告 t_1。容易看出这一点，因为可以改变从报告 t_i 开始的一种安排，其中 $i>1$，使得报告 t_1 替换报告 t_i。为了看出这一点，注意到因为 $e_1\leqslant e_i$，在接着报告 t_i 之后安排的所有报告都还可以被安排。

一旦包含了报告 t_1，安排尽可能多的报告的调度，就归结为安排尽可能多的在时间 e_1 或 e_1 之后开始的报告的调度。所以，如果已经安排了尽可能多的报告，那么对于除报告 t_1 之外的报告的调度，就是对于在报告 t_1 结束后才开始的原始报告的最优调度。因为当贪心算法建立这个调度时安排了 k 个报告，所以可以应用归纳假设来得出贪心算法已经安排了最多可能的报告。结论是当贪心算法产生具有 $k+1$ 个报告的调度时，贪心算法已经安排了最多可能的报告，即 $k+1$ 个，所以 $P(k+1)$为真。这样就完成归纳步骤，证明对于所有可能的整数 n 来说，$P(n)$为真，也完成了最优性的证明。 ■

3.3.4 强归纳法

存在另一种形式的数学归纳法，它在证明里常常很有用。在这种形式里，使用与前面同样的基础步骤，但是使用不同的归纳步骤。假定对 $j=1, \cdots, k$ 来说 $P(j)$ 为真，证明在这个假定的基础之上 $P(k+1)$ 也必然为真。这称为**强归纳法**(有时也称为**数学归纳法的第二原理**)。

总结一下用来证明对所有正整数 n 来说 $P(n)$ 为真的两个步骤：

基础步骤：证明命题 $P(1)$ 为真。

归纳步骤：证明对每个正整数 k 来说，$[P(1) \wedge P(2) \wedge \cdots \wedge P(k)] \rightarrow P(k+1)$ 为真。

数学归纳法的这两种形式是等价的；即假定其中一种是有效的证明技术，可以证明另外一种也是。把这一点留作练习让读者去证明它。现在给出 3 个例子说明如何使用强归纳法。

例 13 考虑一种游戏，其中两名选手轮流从两堆火柴中的一堆取出任意数目的火柴。取走最后一根火柴的选手获胜。证明如果开始时两堆火柴的数目相同，则第二名选手总是可以保证获胜。

解 设 n 是每堆火柴的数目。将用强归纳法来证明 $P(n)$，即命题：当每堆开始有 n 根火柴时，第二名选手可以获胜。

基础步骤：当 $n=1$ 时，第一名选手只有一种选择，从某一堆中取走一根火柴，剩下一堆只有一根，第二名选手可以取走这根火柴而获胜。

归纳步骤：假设对于所有满足 $1 \leqslant j \leqslant k$ 的 j 来说 $P(j)$ 为真，也就是说，每当在游戏开始时两堆各有 j 根火柴，其中 $1 \leqslant j \leqslant k$，第二名选手就总是可以获胜。现在假设在游戏开始时两堆各有 $k+1$ 根火柴，并且假设第一名选手从某一堆取走 j 根火柴($1 \leqslant j \leqslant k$)，在该堆留下 $k+1-j$ 根火柴。第二名选手从另一堆取走同样数目的火柴使得两堆各有 $k+1-j$ 根火柴。因为 $1 \leqslant k+1-j \leqslant k$，根据归纳假设，第二名选手总是可以获胜。注意到如果第一名选手从某一堆取走所有 $k+1$ 根火柴，则第二名选手通过取走所有剩下的火柴就能获胜，这样就完成了证明。 ■

例 14 证明：若 n 是大于 1 的整数，则 n 可以写成素数之积。

解 设 $P(n)$ 是命题：n 可以写成素数之积。

基础步骤：$P(2)$ 为真，因为 2 可以写成一个素数之积，即它自身。(注意 $P(2)$ 是需要证明的第一个情形。)

归纳步骤：假定对所有满足 $j \leqslant k$ 的正整数 j 来说 $P(j)$ 为真。要完成归纳步骤，就必须证明在这个假定下 $P(k+1)$ 为真。

有两种要考虑的情形，即当 $k+1$ 是素数时和当 $k+1$ 是合数时。若 $k+1$ 是素数，则立即看出 $P(k+1)$ 为真。否则，$k+1$ 是合数并且可以写成满足 $2 \leqslant a \leqslant b < k+1$ 的两个整数 a 和 b 之积。根据归纳假设，a 和 b 都可以写成素数之积。因此，若 $k+1$ 是合数，则它可以写成素数之积，即在 a 的因子分解中的那些素数与在 b 的因子分解中的那些素数之积。 ■

注意 因为 1 是素数之积，即不包含任何素数的空积，所以可以在例 14 里用 $P(1)$ 作为基础步骤来开始证明。没有选择这样做是因为许多人对此感到迷惑不解。

注意例 14 完成了对算术基本定理的证明，该定理断言：每个非负整数可以唯一地写成以非降顺序排列的素数之积。在 2.6 节里证明过整数最多有一种这样的素因子分解。例 14 证明至少有一种这样的分解。

用数学归纳法原理来代替强归纳法去证明例 14 里的结果是很难的。不过，例 15 说明用数学归纳法原理或强归纳法都容易证明某些结果。

例 15 证明：仅用 4 分和 5 分邮票就可以组成超过或等于 12 分的每种邮资。

解 将要用数学归纳法原理来证明这个结果。然后用强归纳法证明。设 $P(n)$ 是命题：可以用 4 分和 5 分邮票来组成 n 分邮资。首先使用数学归纳法原理。

基础步骤：可以用 3 个 4 分邮票来组成 12 分邮资。

归纳步骤：假定 $P(k)$ 为真，所以可以用 4 分和 5 分邮票来组成 k 分邮资。若至少用了一个 4 分邮票，则用一个 5 分邮票代替它，就组成 $k+1$ 分邮资。若没有用任何 4 分邮票，则仅用了 5 分的邮票来组成 k 分邮资。因为 $k \geqslant 12$，所以至少用了 3 个 5 分邮票。所以，用 4 个 4 分邮票来代替 3 个 5 分邮票，就组成了 $k+1$ 分邮资。这完成了归纳步骤以及根据数学归纳法原理的证明。

其次，将要使用强归纳法。将要证明可以组成 12、13、14 和 15 分邮资，然后证明如何对 $k \geqslant 15$ 来说从 $k-3$ 分邮资得出 $k+1$ 分邮资。

基础步骤：可以分别用 3 个 4 分邮票、2 个 4 分邮票和 1 个 5 分邮票、1 个 4 分邮票和 2 个 5 分邮票，以及 3 个 5 分邮票，来组成 12、13、14 和 15 分邮资。

归纳步骤：设 $k \geqslant 15$。假定可以组成 j 分邮资，其中 $12 \leqslant j \leqslant k$。为了组成 $k+1$ 分邮资，用组成 $k-3$ 分邮资的邮票加上 1 个 4 分邮票。这完成了归纳步骤以及根据数学归纳法第二原理的证明。

（除了这里描述的方法以外，还有处理这个问题的其他方法。读者能否找出不使用数学归纳法的解答？） ■

注意 例 15 说明如何运用强归纳法处理某些情形，其中仅对充分大的 n 值来说归纳步骤才是有效的。具体说来，为了证明对 $n=j, j+1, j+2, \cdots$ 来说 $P(n)$ 为真，其中 j 是整数，首先证明 $P(j)$，$P(j+1)$，$P(j+2)$，…，$P(l)$ 都为真（基础步骤），然后证明对每个整数 $k \geqslant l$ 来说，$[P(j) \wedge P(j+1) \wedge P(j+2) \wedge \cdots \wedge P(k)] \rightarrow P(k+1)$ 为真（归纳步骤）。例如，例 15 解答里的第二个证明的基础步骤证明 $P(12)$、$P(13)$、$P(14)$ 和 $P(15)$ 都为真。需要分别地证明这些情形，因为归纳步骤证明 $[P(12) \wedge P(13) \wedge \cdots \wedge P(k)] \rightarrow P(k+1)$，它仅当 $k \geqslant 15$ 时才成立。

3.3.5 良序性

数学归纳法的有效性来源于下面这个关于整数集的基本公理。

良序性 每个非空的非负整数集都有最小元。

良序性经常可以直接地用在证明里。

例 16 用良序性证明整除算法。回忆一下，整除算法说：若 a 是整数而且 d 是正整数，则存在唯一的整数 q 和 r 满足 $0 \leqslant r < d$ 和 $a = dq + r$。

解 设 S 是形如 $a-dq$ 的非负整数的集合，其中 q 是整数。这个集合非空，因为 $-dq$ 可以任意大（取 q 是绝对值很大的负整数）。根据良序性，S 有最小元 $r = a - dq_0$。

整数 r 非负而且 $r < d$。若不是这样，则 S 里存在更小的非负整数，即 $a - d(q_0+1)$。为了看出这一点，假设 $r \geqslant d$。因为 $a = dq_0 + r$，所以 $a - d(q_0+1) = a - dq_0 - d = r - d \geqslant 0$。因此，存在满足 $0 \leqslant r < d$ 的整数 r 和 q。证明 q 和 r 都是唯一的，留给读者作为练习。 ■

例 17 在一种主客场循环赛中，每个选手与其他每个选手恰好比赛一次并且每次比赛分出胜负。所谓选手 $p_1, p_2, \cdots, p_m$ 形成回路，就是 p_1 战胜 p_2，p_2 战胜 p_3，…，p_{m-1} 战胜 p_m，p_m 战胜 p_1。用良序原理证明：如果在主客场循环赛的选手中存在长度为 m 的回路（$m \geqslant 3$），则必定存在这些选手中三个选手的回路。

解 假设不存在三个选手的回路。因为在主客场循环赛中至少有一个回路，所以存在长度为 n 的回路的所有正整数 n 的集合是非空的。根据良序性，这个正整数集合有最小元 k，根据假设 k 必定大于 3。因此，存在选手回路 p_1，p_2，…，p_k 并且不存在更短的回路。

现在假设不存在这些选手中三个选手的回路，所以 $k>3$。考虑这个回路的前三个元素 p_1，p_2，p_3。在 p_1 与 p_3 之间的比赛有两种可能的结果。如果 p_3 战胜 p_1，那么 p_1，p_2，p_3 就是长度为 3 的回路，与不存在三个选手的回路的假设相矛盾。因此，必定是 p_1 战胜 p_3。这意味着可以从回路 p_1，p_2，p_3，…，p_k 中忽略 p_2 来获得长度为 $k-1$ 的回路 p_1，p_3，p_4，…，p_k，与最短回路长度为 k 的假设相矛盾。结论是必定存在长度为 3 的回路。 ■

无穷下降法 *无穷下降法*，这种证明方法是皮埃尔·德·费马在 17 世纪引入的。无穷下降法通常用来证明命题函数 $P(n)$，对于所有正整数 k 来说，$P(k)$都为假。这个方法基于这样的事实：如果对于至少一个整数 k 来说，$P(k)$为真，那么良序性就蕴含着存在最小的正整数 s，使得 $P(s)$为真。这个方法继续寻找满足 $s'<s$ 的正整数 s'，使得对于 s'来说 $P(s')$为真。所以对于所有正整数来说，$P(n)$必定都为假。(这种技术被称为*无穷下降法*，是因为寻找使得命题函数为真的更小的正整数的过程可以无穷地继续下去，产生一个递减的正整数的无穷序列，根据良序性这是不可能的。)无穷下降法通常用来证明某些方程没有整数解。具体说来，费马用它证明了费马大定理 $n=4$ 的情形，这种情形说方程 $x^4+y^4=z^4$ 没有正整数解。在例 18 中解释无穷下降法的用法。

例 18 在 1.5 节例 21 中证明了$\sqrt{2}$是无理数。这里将用无穷下降法给出这个事实的不同证明。首先，假设$\sqrt{2}$是有理数。于是存在正整数 m 和 n 使得 $\sqrt{2}=m/n$。根据良序性，存在最小正整数 N，使得对于某个正整数 M 来说 $\sqrt{2}=M/N$。(这使得 N 是等于$\sqrt{2}$的两个正整数之商的最小可能的分母。)

为了用无穷下降法完成证明，将证明 $\sqrt{2}=(2N-M)/(M-N)$并且 $0<M-N<N$。这与选择 N 是对于某个正整数 M 来说使得 $\sqrt{2}=M/N$ 的最小正整数相矛盾。为了要证明$\sqrt{2}=(2N-M)/(M-N)$，只需证明$(2N-M)/(M-N)=M/N$。注意到因为$(M/N)^2=2$，所以 $M^2=2N^2$。因此

$$\frac{2N-M}{M-N}=\frac{(2N-M)N}{(M-N)N}=\frac{2N^2-MN}{(M-N)N}=\frac{M^2-MN}{(M-N)N}=\frac{(M-N)M}{(M-N)N}=\frac{M}{N}$$

为了完成证明，只需要证明分母 $M-N$ 是正的并且小于 N。为了看出这一点，注意到因为$1<\sqrt{2}<2$ 并且$\sqrt{2}=M/N$，所以 $1<M/N<2$，因此 $N<M<2N$。减去 N，就得到 $0<M-N<N$。 ■

3.3.6 为什么数学归纳法有效

为什么数学归纳法是一种有效的证明技术？原因来源于良序性。假定知道 $P(1)$为真，而且对所有正整数 k 来说，命题 $P(k)\rightarrow P(k+1)$为真。为了证明对所有正整数来说 $P(n)$都为真，假定至少存在一个使 $P(n)$为假的正整数，那么使 $P(n)$为假的正整数集合 S 非空。因此，根据良序性，S 有一个最小元，把它表示成 m。可以知道 m 不是 1，因为 $P(1)$为真。因为 m 是正的而且大于 1，所以 $m-1$ 是一个正整数。另外，因为 $m-1$ 小于 m，它不属于 S，所以 $P(m-1)$必然为真。因为蕴含式 $P(m-1)\rightarrow P(m)$也为真，所以实际情形必然是 $P(m)$为真。这与对 m 的选择相矛盾。因此，对每个正整数 n 来说，$P(n)$必然为真。

练习

1. 找出前 n 个正偶数之和的公式。
2. 用数学归纳法证明在练习 1 里找出的公式。

3. 用数学归纳法证明:每当 n 是非负整数时,就有 $3+3\cdot5+3\cdot5^2+\cdots+3\cdot5^n=3(5^{n+1}-1)/4$。
4. 用数学归纳法证明:每当 n 是非负整数时,就有 $2-2\cdot7+2\cdot7^2-\cdots+2(-7)^n=(1-(-7)^{n+1})/4$。
5. 通过对小的 n 值检查表达式

$$\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots+\frac{1}{2^n}$$

的值,来找出这个表达式的公式。用数学归纳法证明你的结果。
6. 通过对小的 n 值检查表达式

$$\frac{1}{1\cdot2}+\frac{1}{2\cdot3}+\cdots+\frac{1}{n(n+1)}$$

的值,来找出这个表达式的公式。用数学归纳法证明你的结果。
7. 证明:每当 n 是正整数时,就有 $1^2+2^2+\cdots+n^2=n(n+1)(2n+1)/6$。
8. 证明:每当 n 是正整数时,就有 $1^3+2^3+\cdots+n^3=[n(n+1)/2]^2$。
9. 证明:每当 n 是非负整数时,就有 $1^2+3^2+5^2+\cdots+(2n+1)^2=(n+1)(2n+1)(2n+3)/3$。
10. 证明:每当 n 是正整数时,就有 $1\cdot1!+2\cdot2!+\cdots+n\cdot n!=(n+1)!-1$
*11. 用数学归纳法证明:若 $h>-1$,则对所有非负整数 n 来说,$1+nh\leqslant(1+h)^n$。这称为伯努利不等式。
12. 证明:每当 n 是大于6的正整数时,就有 $3^n<n!$。
13. 证明:每当 n 是大于4的正整数时,就有 $2^n>n^2$。
14. 用数学归纳法证明:每当 n 是大于1的正整数时,就有 $n!<n^n$。
15. 用数学归纳法证明:每当 n 是正整数时,就有 $1\cdot2+2\cdot3+\cdots+n(n+1)=n(n+1)(n+2)/3$。
16. 用数学归纳法证明:$1\cdot2\cdot3+2\cdot3\cdot4+\cdots+n(n+1)(n+2)=n(n+1)(n+2)(n+3)/4$。
17. 证明:每当 n 是正整数时,就有 $1^2-2^2+3^2-\cdots+(-1)^{n-1}n^2=(-1)^{n-1}n(n+1)/2$。
18. 证明:每当 n 是大于1的正整数时,就有

$$1+\frac{1}{4}+\frac{1}{9}+\cdots+\frac{1}{n^2}<2-\frac{1}{n}$$

19. 证明:可以仅用3分邮票和5分邮票,来组成大于7分的正整数分值的任何邮资。
20. 用数学归纳法证明:每当 n 是非负整数时,就有3整除 n^3+2n。
21. 用数学归纳法证明:每当 n 是非负整数时,就有5整除 n^5-n。
22. 用数学归纳法证明:每当 n 是非负整数时,就有6整除 n^3-n。
*23. 用数学归纳法证明:每当 n 是正奇数时,就有 n^2-1 被8整除。
24. 用数学归纳法证明:每当 n 是大于3的整数时,就有 $n^2-7n+12$ 是非负的。
25. 用数学归纳法证明:每当 n 是大于或等于2的整数时,一个带有 n 个元素的集合就有 $n(n-1)/2$ 个恰好包含两个元素的子集合。
*26. 用数学归纳法证明:每当 n 是大于或等于3的整数时,一个带有 n 个元素的集合就有 $n(n-1)(n-2)/6$ 个恰好包含三个元素的子集合。
27. 用数学归纳法证明:每当 n 是正整数时,就有 $\sum_{j=1}^{n} j^4=n(n+1)(2n+1)(3n^2+3n-1)/30$。
28. 对哪些非负整数 n 来说,有 $n^2\leqslant n!$?用数学归纳法证明你的答案。
29. 对哪些非负整数 n 来说,有 $2n+3\leqslant2^n$?用数学归纳法证明你的答案。
30. 用数学归纳法证明:每当 n 是正整数时,就有 $1/(2n)\leqslant[1\cdot3\cdot5\cdot\cdots\cdot(2n-1)]/(2\cdot4\cdot\cdots\cdot2n)$。
31. a) 确定仅用5分邮票和6分邮票可以组成哪些数量的邮资?
 b) 用数学归纳法证明你对(a)的答案。
 c) 用数学归纳法第二原理证明你对(a)的答案。
32. 仅用10分邮票和25分邮票可以组成哪些数量的邮资?
33. 一种自动柜员机中只有20美元和50美元面额的钞票。假定这种机器中两种美元的供应量是无限的,这种机器可以分发哪些数量的美元?用数学归纳法的一种形式证明你的答案。

34. 假设一种巧克力棒由排列成长方形的 n 个方块组成。整个棒或棒的较小的长方形块可以沿着分隔方块的垂直线或水平线折断。假设一次只能折断一块，确定为了把整个棒折断成 n 个分开的方块，必须先后折断多少次。用强归纳法证明你的答案。
35. 考虑拿子游戏的如下变种。这个游戏从 n 根火柴开始。两名选手轮流取走火柴，每次取一根、两根或三根。取走最后一根火柴的选手落败。用强归纳法证明：如果每名选手按照最好可能的策略来游戏，那么若对于某个非负整数 j 来说，$n = 4j$、$4j+2$ 或 $4j+3$ 时，则第一名选手获胜，而在 $n = 4j+1$ 的其他情形下第二名选手获胜。
36. 用数学归纳法证明：$\sum_{k=1}^{n} k2^k = (n-1)2^{n+1} + 2$。
37. 证明：如果 n 是正整数，有 $\sum_{\{a_1,\cdots,a_k\}\subseteq\{1,2,\cdots,n\}} \frac{1}{a_1 a_2 \cdots a_k} = n$。（在这里求和是在 n 个最小正整数的集合的所有非空子集上进行。）
38. 用数学归纳法证明：给定有 $n+1$ 个正整数的集合，这些正整数都不超过 $2n$，在这个集合里至少存在一个整数，它整除在这个集合里的另外一个整数。
*39. 棋盘上的马可以（沿任意方向）水平地移动一格并（沿任意方向）垂直地移动两格，或者（沿任意方向）水平地移动两格并（沿任意方向）垂直地移动一格。用数学归纳法证明：在包含所有方格 (m,n) 的无穷棋盘上，对于每个方格来说，马可以从 $(0,0)$ 角落开始，通过有穷的移动序列来访问这个方格，其中 m 和 n 都是非负整数。（提示：对变量 $s = m + n$ 用归纳法。）
40. 假设从一堆 n 块石头开始，通过连续地把一堆石头分成较小的两堆石头，把开始的这堆石头分成 n 堆，每堆只有 1 块石头。每次分开一堆石头时，就把所分出的较小的两堆石头的数目相乘，即如果分出的这两堆分别有 r 和 s 块石头，则计算出 rs。证明：无论如何分这些堆，每一步所计算出来的乘积之和等于 $n(n-1)/2$。
41. （本题需要微积分知识）用数学归纳法证明：每当 n 是正整数时，就有 $f(x) = x^n$ 的导数等于 nx^{n-1}。（对归纳步骤来说，使用导数的乘法规则。）
42. 假定

$$\mathbf{A} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

其中 a 和 b 都是正实数。证明：对每个正整数 n 来说

$$\mathbf{A}^n = \begin{bmatrix} a^n & 0 \\ 0 & b^n \end{bmatrix}$$

43. 假定 $\mathbf{A}$ 和 $\mathbf{B}$ 都是方阵且具有性质 $\mathbf{AB} = \mathbf{BA}$。证明：对每个正整数 n 来说，有 $\mathbf{AB}^n = \mathbf{B}^n\mathbf{A}$。
44. 假定 m 是正整数。用数学归纳法证明：若 a 和 b 都是整数且 $a \equiv b(\mathrm{mod}\, m)$，则每当 k 是非负整数时，就有 $a^k \equiv b^k(\mathrm{mod}\, m)$。
45. 用数学归纳法证明：若 $A_1, A_2, \cdots, A_n$ 和 B 都是集合，则

$$(A_1 \cup A_2 \cup \cdots \cup A_n) \cap B$$
$$= (A_1 \cap B) \cup (A_2 \cap B) \cup \cdots \cup (A_n \cap B)$$

46. 证明：若 $A_1, A_2, \cdots, A_n$ 和 $B_1, B_2, \cdots, B_n$ 都是集合，使得对 $k = 1, 2, \cdots, n$ 来说 $A_k \subseteq B_k$，则

a) $\bigcup_{k=1}^{n} A_k \subseteq \bigcup_{k=1}^{n} B_k$　　b) $\bigcap_{k=1}^{n} A_k \subseteq \bigcap_{k=1}^{n} B_k$

47. 用数学归纳法证明：若 $A_1, A_2, \cdots, A_n$ 都是论域 U 的子集，则

$$\overline{\bigcup_{k=1}^{n} A_k} = \bigcap_{k=1}^{n} \overline{A_k}$$

48. 用数学归纳法证明：每当 $p_1, p_2, \cdots, p_n$ 都是命题时，$\neg(p_1 \vee p_2 \vee \cdots \vee p_n)$ 就等价于 $\neg p_1 \wedge \neg p_2 \wedge \cdots \wedge \neg p_n$。

*49. 证明：每当 $p_1, p_2, \cdots, p_n$ 都是命题时，

$[(p_1 \rightarrow p_2) \wedge (p_2 \rightarrow p_3) \wedge \cdots \wedge (p_{n-1} \rightarrow p_n)] \rightarrow [(p_1 \wedge p_2 \wedge \cdots \wedge p_{n-1}) \rightarrow p_n]$

就是重言式。

50. 下面这个“证明”错在哪里？

定理：对于每个正整数 n 来说，$\sum_{i=1}^{n} i = \left(n+\frac{1}{2}\right)^2/2$。

基础步骤：对于 $n=1$ 来说这个公式为真。

归纳步骤：假设 $\sum_{i=1}^{n} i = \left(n+\frac{1}{2}\right)^2/2$。于是 $\sum_{i=1}^{n+1} i = \left(\sum_{i=1}^{n} i\right) + (n+1)$。根据归纳假设，

$\sum_{i=1}^{n+1} i = \left(n+\frac{1}{2}\right)^2/2 + n + 1 = \left(n^2+n+\frac{1}{4}\right)/2 + n + 1 = \left(n^2+3n+\frac{9}{4}\right)/2 = \left(n+\frac{3}{2}\right)^2/2 =$

$\left[(n+1)+\frac{1}{2}\right]^2/2$，完成了归纳步骤。

51. 下述的关于所有马都是相同颜色的“证明”错在哪里？

设 $P(n)$ 是命题：在 n 匹马的集合里所有马都是相同颜色的。

基础步骤：显然，$P(1)$ 为真。

归纳步骤：现在假定 $P(k)$ 为真，所以在任意 k 匹马的集合里，所有马都是相同颜色的。考虑任意的 $k+1$ 匹马；把这些马编号成 $1,2,3,\cdots,k+1$。现在这些马的前 k 匹必然都有相同颜色，而且这些马的后 k 匹也必然都有相同颜色。因为前 k 匹马的集合与后 k 匹马的集合重叠，所以所有 $k+1$ 匹马必然都是相同颜色。这就证明了 $P(k+1)$ 为真并且完成了归纳证明。

52. 下面的“证明”错在哪里？

定理：对于每个正整数 n 来说，如果 x 和 y 都是满足 $\max(x,y)=n$ 的正整数，则 $x=y$。

基础步骤：假设 $n=1$。如果 $\max(x,y)=1$ 并且 x 和 y 都是正整数，就有 $x=1$ 和 $y=1$。

归纳步骤：设 k 是正整数。假设每当 $\max(x,y)=k$ 并且 x 和 y 都是正整数时，则 $x=y$。现在设 $\max(x,y)=k+1$，其中 x 和 y 都是正整数。则 $\max(x-1,y-1)=k$，所以根据归纳假设，$x-1=y-1$。所以 $x=y$，完成了归纳步骤。

53. 下面利用强归纳法的“证明”错在哪里？

定理：对于每个非负整数 n 来说，$5n=0$。

基础步骤：$5 \cdot 0 = 0$。

归纳步骤：假设对于满足 $0 \leqslant j \leqslant k$ 的所有非负整数 j 来说，$5j=0$。写 $k+1=i+j$，其中 i 和 j 都是小于 $k+1$ 的自然数。根据归纳假设，$5(k+1)=5(i+j)=5i+5j=0+0=0$。

*54. 找出下列“证明”的错误：每当 a 是非零实数时，对所有非负整数 n 来说，有 $a^n=1$。

基础步骤：根据 a^0 的定义，$a^0=1$ 为真。

归纳步骤：假定对满足 $j \leqslant k$ 的所有非负整数 j 来说，$a^j=1$。则注意 $a^{k+1} = \frac{a^k \cdot a^k}{a^{k-1}} = \frac{1 \cdot 1}{1} = 1$

*55. 通过证明从良序性得出强归纳法，来证明强归纳法是有效的证明方法。

*56. 证明：为了证明对所有正整数 n 来说，$P(n)$ 为真，数学归纳法的下列形式是有效的证明方法。

基础步骤：$P(1)$ 和 $P(2)$ 都为真。

归纳步骤：对每个正整数 k 来说，若 $P(k)$ 和 $P(k+1)$ 都为真，则 $P(k+2)$ 为真。

在练习 57 和练习 58 中，H_n 表示第 n 个调和数。

*57. 用数学归纳法证明：每当 n 是非负整数时，就有 $H_{2^n} \leqslant 1+n$。

*58. 用数学归纳法证明：$H_1 + H_2 + \cdots + H_n = (n+1)H_n - n$。

*59. 证明：$1+\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{3}}+\cdots+\frac{1}{\sqrt{n}} > 2(\sqrt{n+1}-1)$。

*60. 证明：若 n 条直线中任何两条都不平行，任何三条都不共点，则这些直线把平面分割成 $(n^2+n+2)/2$ 个区域。

**61. 设 $a_1, a_2, \cdots, a_n$ 都是正实数。这些数的算术均值定义成 $A=(a_1+a_2+\cdots+a_n)/n$，而这些数的几何均值定义成 $G=(a_1a_2\cdots a_n)^{1/n}$。用数学归纳法证明：$A \geqslant G$。

*62. 用数学归纳法证明：当 n 是正整数时，就有 21 整除 $4^{n+1}+5^{2n-1}$。

63. 用数学归纳法证明 2.6 节引理 2，该引理说：若 p 是素数且 $p \mid a_1a_2\cdots a_n$，其中对 $i=1,2,\cdots,n$ 来说 a_i 是整数，则对某个整数 i 来说 $p \mid a_i$。

64. 用无穷下降法证明方程 $8x^4+4y^4+2z^4=w^4$ 没有 x、y、z 和 w 的正整数解。

65. 用无穷下降法证明 $w^2+x^2+y^2+z^2=2wxyz$ 没有 x、y、z 和 w 的正整数解。(提示：首先证明如果这个方程成立，那么 x、y、z 和 w 全都是偶数。然后证明这四个整数全都必须被 4 整除，被 8 整除，等等。)

*66. 良序性可以用来证明：两个正整数有唯一的最大公因子。设 a 和 b 都是正整数，设 S 是形如 $as+bt$ 的正整数的集合，其中 s 和 t 都是正整数。

a) 证明：S 非空。

b) 用良序性证明：S 有最小元 c。

c) 证明：若 d 是 a 和 b 的公因子，则 d 是 c 的因子。

d) 证明：$c \mid a$ 和 $c \mid b$。(提示：首先假定 $c \nmid a$。则 $a=qc+r$，其中 $0<r<c$。证明 $r \in S$，这与对 c 的选择相矛盾。)

e) 从(c)和(d)得出：a 和 b 的最大公因子存在。通过证明两个正整数的最大公因子是唯一的，来完成证明。

*67. 证明：若 $a_1, a_2, \cdots, a_n$ 是 n 个不同的实数，则无论在它们的乘积中插入多少对括号，计算这 n 个数之积都要使用 $n-1$ 次乘法。(提示：利用强归纳法并且考虑最后一次的乘法。)

68. 用 L 形状的碎片去铺满去掉左上角格子的 4×4 棋盘。

69. 用 L 形状的碎片去铺满去掉左上角格子的 8×8 棋盘。

70. 证明或反驳：每当 n 是正整数时，就可以用 L 形状的碎片完全地覆盖下面形状的所有棋盘。

a) 3×2^n　　b) 6×2^n　　c) $3^n \times 3^n$　　d) $6^n \times 6^n$

*71. 证明：用去掉了一个 $1\times1\times1$ 立方体的 $2\times2\times2$ 立方体，可以完全地覆盖去掉了一个 $1\times1\times1$ 立方体的三维 $2^n \times 2^n \times 2^n$ 棋盘。

*72. 证明：若 $n>5$，n 是奇数，且 n 不能被 3 整除，则用 L 形状的碎片可以完全地覆盖去掉了一个格子的 $n \times n$ 棋盘。

73. 证明：可以用 L 形状的碎片来铺满去掉了一个角上格子的 5×5 棋盘。

*74. 找出不能用 L 形状的碎片来铺满的去掉了一个格子的 5×5 棋盘。证明这种铺法铺不出这个棋盘来。

75. 设 a 是整数而 d 是正整数。证明：在例 16 中论证过其存在性的、满足 $a=dq+r$ 和 $0 \leqslant r<d$ 的整数 q 和 r 是唯一的。

☞76. 用数学归纳法原理证明：若 $P(b)$ 为真，并且对满足 $k \geqslant b$ 的所有正整数 k 来说，蕴含式 $P(k) \rightarrow P(k+1)$ 为真，则对 $n=b, b+1, b+2, \cdots$ 来说 $P(n)$ 为真，其中 b 是整数。

**77. 你能用良序性证明下面的命题吗?"每一个正整数都可以用不超过 15 个英语单词来描述"。

78. 用良序性证明：若 x 和 y 是满足 $x<y$ 的实数，则存在有理数 r 满足 $x<r<y$。(提示：证明存在正整数 A 满足 $A>1/(y-x)$。然后通过考虑数 $\lfloor x \rfloor + j/A$，其中 j 是正整数，来证明存在有理数 r 具有介于 x 和 y 之间的分母 A。)

3.4 递归定义与结构归纳法

3.4.1 引言

有时难以用明确的方式来定义一个对象。不过，用这个对象来定义它自身，这也许是容易的。这种过程称为*递归*。例如，图 3-11 所示的图画是递归地产生的。首先，给出一幅原图。然

后实现在前一幅图画的中央递归地放上更小的图画这样一个过程。

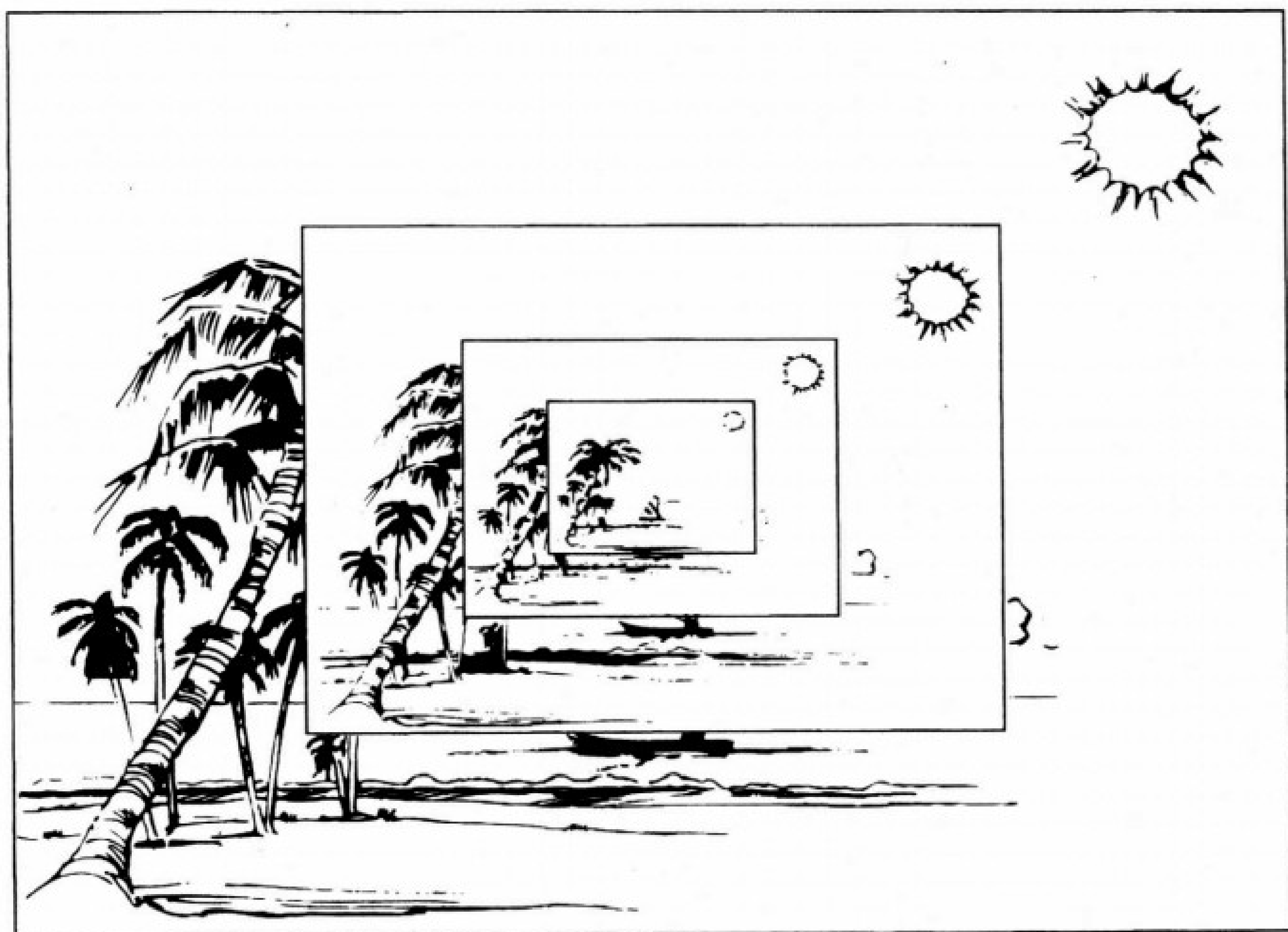


图 3-11 递归定义的图画

可以用递归来定义序列、函数和集合。在前面的讨论里用显式的公式来规定序列里的项。例如，用“对 $n=0, 1, 2, \cdots$来说，$a_n=2^n$”来给出 2 的幂的序列。不过，通过给出这个序列的第一项，即 $a_0=1$，以及从该序列前面的项来求出当前项的公式，即对 $n=0, 1, 2, \cdots$来说 $a_{n+1}=2a_n$，也可以定义这个序列。当通过规定如何从前面的各项求出序列各项来递归地定义序列时，可以用归纳法来证明关于这个序列的结果。

当递归地定义集合时，在基础步骤里规定一些初始元素，并且在递归步骤里提供一条规则，从已有的那些元素来构造新的元素。为了证明关于递归地定义的集合的结果，使用所谓的结构归纳法。

3.4.2 递归地定义函数

为了定义以非负整数集合作为其定义域的函数，使用两个步骤：

基础步骤：规定这个函数在 0 处的值。

递归步骤：给出从较小的整数处的值来求出当前的值的规则。

这样的定义称为递归定义或归纳定义。



例 1 假定 f 是用

$$f(0)=3$$
$$f(n+1)=2f(n)+3$$

来递归地定义的。求出 $f(1)$、$f(2)$、$f(3)$和 $f(4)$。

解 从这个递归定义得出

$$f(1) = 2f(0) + 3 = 2 \cdot 3 + 3 = 9$$
$$f(2) = 2f(1) + 3 = 2 \cdot 9 + 3 = 21$$
$$f(3) = 2f(2) + 3 = 2 \cdot 21 + 3 = 45$$
$$f(4) = 2f(3) + 3 = 2 \cdot 45 + 3 = 93$$
■

许多函数都可以利用它们的递归定义来研究。阶乘函数就是一个这样的例子。

例 2 给出阶乘函数 $F(n)=n!$ 的归纳定义。

解 可以通过规定阶乘函数的初值，即 $F(0)=1$，并且给出从 $F(n)$求出 $F(n+1)$的规则，来定义这个函数。要得出这个结果，注意通过乘以 $n+1$ 就从 $n!$ 计算出$(n+1)!$。因此，所需要的规则是 $F(n+1)=(n+1)F(n)$。 ■

为了从在例 2 里求出的递归定义来确定阶乘函数的一个值，比如 $F(5)=5!$，有必要多次使用 $F(n)$表示 $F(n+1)$的规则：

$$F(5) = 5F(4) = 5 \cdot 4F(3) = 5 \cdot 4 \cdot 3F(2) = 5 \cdot 4 \cdot 3 \cdot 2F(1)$$
$$= 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 \cdot F(0) = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 \cdot 1 = 120$$

一旦 $F(0)$是出现的唯一函数值，就不需要任何更多的归约。剩下要做的唯一事情是把 $F(0)$的值插入到公式里。

递归定义的函数是良定义的。这是数学归纳法原理的一个结果(见本节末的练习 56)。在下面的例子里给出递归定义的其他例子。

例 3 给出 a^n 的递归定义，其中 a 是非零实数而且 n 是非负整数。

解 这个递归定义包括两个部分。首先规定 a^0，即 $a^0=1$。然后给出从 a^n 求出 a^{n+1} 的规则，即对$n=0, 1, 2, 3, \cdots$，$a^{n+1}=a \cdot a^n$。这两个等式对所有非负整数唯一地定义了 a^n。 ■

例 4 给出 $\sum_{k=0}^{n} a_k$ 的递归定义。

解 这个递归定义的第一部分是

$$\sum_{k=0}^{0} a_k = a_0$$

第二部分是

$$\sum_{k=0}^{n+1} a_k = \left(\sum_{k=0}^{n} a_k\right) + a_{n+1}$$
■

在函数的某些递归定义里，规定了函数在前 k 个正整数处的值，而且给出了从一个较大的整数之前的部分或全部 k 个整数处的函数值来确定在该整数处的函数值的规则。从强归纳法可以得出结论说这样的定义产生良定义的函数(见本节末的练习 57)。

定义 1 斐波那契[⊖]数 $f_0, f_1, f_2, \cdots$是用方程组 $f_0=0$，$f_1=1$ 和 $f_n=f_{n-1}+f_{n-2}$ 来定义的，其中 $n=2, 3, 4, \cdots$。

例 5 求斐波那契数 f_2, f_3, f_4, f_5, f_6。

解 因为这个定义的第一部分说 $f_0=0$ 和 $f_1=1$，所以从这个定义的第二部分得出

⊖ 斐波那契(Fibonacci，约 1170—1250)诞生在意大利的商业中心比萨，称为比萨的列奥那多(Leonardo)。斐波那契是一位商人。他遍游中东各地，在那里接触到阿拉伯数学。在他的著作《*Liber Abaci*》中，斐波那契向欧洲人介绍了阿拉伯的数字记号和算术的算法。著名的兔子问题(在 6.1 节描述)就出现在此书。斐波那契还写过关于几何学和三角学以及关于丢番图方程的各种论著，丢番图方程是关于寻找方程的整数解的。

$$f_2 = f_1 + f_0 = 1+0 = 1$$
$$f_3 = f_2 + f_1 = 1+1 = 2$$
$$f_4 = f_3 + f_2 = 2+1 = 3$$
$$f_5 = f_4 + f_3 = 3+2 = 5$$
$$f_6 = f_5 + f_4 = 5+3 = 8$$ ■

可以用斐波那契数的递归定义来证明这些数的许多性质。在例 6 里给出一个这样的性质。

例 6 证明：每当 $n\geqslant 3$ 时，有 $f_n>\alpha^{n-2}$，其中 $\alpha=(1+\sqrt{5})/2$。

解 可以用强归纳法来证明这个不等式。设 $P(n)$ 是命题：$f_n>\alpha^{n-2}$。想要证明的是每当 n 是大于或等于 3 的整数时，有 $P(n)$ 为真。

基础步骤：首先注意到，

$$\alpha < 2 = f_3,\quad \alpha^2 = (3+\sqrt{5})/2 < 3 = f_4$$

所以 $P(3)$ 和 $P(4)$ 都为真。

归纳步骤：现在假定 $P(j)$ 为真，即对所有满足 $3\leqslant j\leqslant k$ 的整数 j 来说有 $f_j>\alpha^{j-2}$，其中 $k\geqslant 4$。必须证明 $P(k+1)$ 为真，即 $f_{k+1}>\alpha^{k-1}$。因为 α 是 $x^2-x-1=0$ 的解(二次方程求根公式说明这一点)，所以得出 $\alpha^2=\alpha+1$。因此，

$$\alpha^{k-1} = \alpha^2\cdot\alpha^{k-3} = (\alpha+1)\alpha^{k-3} = \alpha\cdot\alpha^{k-3}+1\cdot\alpha^{k-3} = \alpha^{k-2}+\alpha^{k-3}$$

根据归纳假设，若 $k\geqslant 4$ 则得出

$$f_{k-1} > \alpha^{k-3}, f_k > \alpha^{k-2}$$

因此就有

$$f_{k+1} = f_k + f_{k-1} > \alpha^{k-2}+\alpha^{k-3} = \alpha^{k-1}$$

由此得出 $P(k+1)$ 为真，证毕。 ■

注意 *归纳步骤证明了每当 $k\geqslant 4$ 时，从对 $3\leqslant j\leqslant k$ 来说 $P(j)$ 为真的假定就得出 $P(k+1)$。因此，归纳步骤没有证明 $P(3)\to P(4)$。因此，不得不单独证明 $P(4)$ 为真。*

现在可以证明：欧几里得算法用 $O(\log b)$ 次除法来求出正整数 a 和 b 的最大公因子，其中 $a\geqslant b$。

定理 1 **拉梅**[⊖]**定理** 设 a 和 b 是满足 $a\geqslant b$ 的正整数。则欧几里得算法为了求出 $\gcd(a, b)$ 而使用的除法的次数小于或等于 b 的十进制位数的 5 倍。

证明 回忆一下，当用欧几里得算法求满足 $a\geqslant b$ 的 $\gcd(a, b)$ 时，得出了下面的等式序列

⊖ **加布里尔·拉梅(Gabriel Lamé, 1795—1870)** 加布里尔·拉梅于 1813 年进入工业高等专科学校，1817 年毕业。他在米内兹高等专科学校继续接受教育，于 1820 年毕业。

1820 年拉梅来到俄国，被任命为圣彼得堡公路与运输学校的校长。在俄国期间他不仅教书，而且设计道路和桥梁。他在 1832 年回到巴黎，帮助成立一家工程公司。不过，他很快离开这个公司，接受了工业高等专科学校的物理学教授职务，担任这个职务直到 1844 年。在此期间，他作为工程顾问活跃在其他领域，担任煤矿的首席工程师并且参与铁路建设。

拉梅对数论、应用数学以及热力学都做出了开创性工作。他最著名的工作包括引入曲线坐标。他研究数论，证明了 $n=7$ 时的费马大定理，以及本文给出的欧几里得算法所用除法次数的上界。

一直以来最伟大的数学家高斯，认为拉梅是法国当时最出色的数学家。不过，法国数学家认为他太实际，而法国科学家认为他太理论化。

（其中$a=r_0$而$b=r_1$）。

$$\begin{aligned} r_0 &= r_1 q_1 + r_2 \quad 0 \leqslant r_2 < r_1 \\ r_1 &= r_2 q_2 + r_3 \quad 0 \leqslant r_3 < r_2 \\ &\vdots \\ r_{n-2} &= r_{n-1} q_{n-1} + r_n \quad 0 \leqslant r_n < r_{n-1} \\ r_{n-1} &= r_n q_n \end{aligned}$$

在这里为了求出 gcd(a, b)而使用了 n 次除法。注意商 q_1, q_2, …, q_{n-1}都至少是 1。另外，$q_n \geqslant 2$,因为 $r_n < r_{n-1}$。这就蕴含着

$$\begin{aligned} r_n &\geqslant 1 = f_2 \\ r_{n-1} &\geqslant 2r_n \geqslant 2f_2 = f_3 \\ r_{n-2} &\geqslant r_{n-1} + r_n \geqslant f_3 + f_2 = f_4 \\ &\vdots \\ r_2 &\geqslant r_3 + r_4 \geqslant f_{n-1} + f_{n-2} = f_n \\ b = r_1 &\geqslant r_2 + r_3 \geqslant f_n + f_{n-1} = f_{n+1} \end{aligned}$$

由此得出，若欧几里得算法为了求出满足 $a \geqslant b$ 的 gcd(a, b)而使用了 n 次除法，则 $b \geqslant f_{n+1}$。从例 6 中知道，对 $n>2$ 来说 $f_{n+1} > \alpha^{n-1}$，其中 $\alpha = (1+\sqrt{5})/2$。因此得出 $b > \alpha^{n-1}$。另外，因为 $\log_{10}\alpha \sim 0.208 > 1/5$，所以可以看出

$$\log_{10} b > (n-1)\log_{10}\alpha > (n-1)/5$$

因此，$n-1 < 5 \cdot \log_{10} b$。现在假定 b 有 k 个十进制位。则 $b < 10^k$ 而且 $\log_{10} b < k$。由此得出 $n-1 < 5k$，而且因为 k 是整数，所以得出 $n \leqslant 5k$。证毕。 □

因为 b 的十进制位数等于$\lfloor \log_{10} b \rfloor + 1$，它小于或等于 $\log_{10} b + 1$，故定理 1 说求出满足 $a \geqslant b$ 的 gcd(a, b)所需要的除法次数小于或等于 $5(\log_{10} b + 1)$。因为 $5(\log_{10} b + 1)$是 $O(\log b)$，故可看出每当 $a > b$ 时，欧几里得算法就用 $O(\log b)$次除法来求出 gcd(a, b)。

3.4.3 递归地定义集合与结构

前面探讨了如何递归地定义函数。现在把注意力转到如何递归地定义集合。正如在函数的递归定义中那样，集合的递归定义有两个部分：*基础步骤*和*递归步骤*。在基础步骤中，规定初始的一些元素。在递归步骤中，给出用来从已知属于集合的元素来构造集合的新元素的规则。递归定义也可以包含一条*排斥规则*，这条规则规定，递归定义的集合仅仅包含基础步骤所规定的以及递归步骤的应用所生成的那些元素。在本书的讨论中，将总是默认排斥规则成立，因而任何元素都不属于递归定义的集合，除非这个元素属于基础步骤所规定的初始的一些元素，或者是可以一次或多次使用递归步骤来生成的。稍后将介绍如何用所谓的结构归纳法技术来证明关于递归定义的集合的结果。

例 7、例 8、例 10 和例 11 解释集合的递归定义。在每个例子中，都说明递归步骤的头几次应用所生成的那些元素。

Extra Examples **例 7** 考虑如下定义的整数集合的子集 S:

基础步骤：$3 \in S$。

递归步骤：若 $x \in S$ 且 $y \in S$，则 $x+y \in S$。

基础步骤中求出的 S 中的新元素是 3，递归步骤的首次应用求出的是 $3+3=6$，递归步骤的第二次应用求出的是 $3+6=6+3=9$ 以及 $6+6=12$，等等。 ■

在对字符串的研究里递归定义起着重要作用(例如，参见第 11 章对形式语言的介绍)。回忆

一下3.2节，字母表Σ上的字符串是Σ里符号的有穷序列。定义2说明可以递归地定义Σ上的字符串的集合Σ^*。

定义2 字母表Σ上的字符串的集合Σ^*递归地定义成：

基础步骤：$\lambda\in\Sigma^*$（其中λ是不包含任何符号的空串）。

递归步骤：若$w\in\Sigma^*$且$x\in\Sigma$时，则$wx\in\Sigma^*$。

字符串的递归定义的基础步骤说空串属于Σ^*。递归步骤说把Σ的符号添加到Σ^*的字符串结尾后面就生成新的字符串。在递归步骤的每次应用中，都生成包含一个更多符号的字符串。

例8 若$\Sigma=\{0,1\}$，则Σ^*中的字符串（即所有位串的集合）是：在基础步骤中规定属于Σ^*的λ，在递归步骤的首次应用中形成的0和1，在递归步骤的第二次应用中形成的00、01、10和11，等等。■

递归定义可用于在递归定义的集合的元素上来定义运算或函数。定义3说明了这一点，定义3是关于两个字符串的连接的，例9是关于字符串的长度的。

定义3 通过连接运算可以组合两个字符串。设Σ是符号的集合，Σ^*是Σ中符号形成的字符串的集合。可以如下定义两个字符串的连接，用·表示：

基础步骤：若$w\in\Sigma^*$，则$w\cdot\lambda=w$，其中λ是空串。

递归步骤：若$w_1\in\Sigma^*$且$w_2\in\Sigma^*$以及$x\in\Sigma$，则$w_1\cdot(w_2x)=(w_1\cdot w_2)x$。

字符串w_1和w_2的连接通常写成w_1w_2，而不是$w_1\cdot w_2$。通过反复应用递归定义，就得出两个字符串w_1和w_2的连接是w_1中的符号后面跟着w_2中的符号。例如，$w_1=abra$和$w_2=cadabra$的连接是$w_1w_2=abracadabra$。

例9　字符串的长度　给出字符串w的长度$l(w)$的递归定义。

解　字符串的长度可以定义成

$$l(\lambda)=0;$$
$$l(wx)=l(w)+1,\text{若 } w\in\Sigma^* \text{ 且 } x\in\Sigma。$$
■

递归定义的另一种重要用途是定义各种类型的合式公式。在例10和例11里说明这一点。

例10　复合命题的合式公式　可以定义关于**T**、**F**、命题变元以及集合$\{\neg,\wedge,\vee,\rightarrow,\leftrightarrow\}$中的运算的复合命题的合式公式的集合。

基础步骤：**T**、**F**和p都是合式公式，其中p是命题变元。

递归步骤：若E和F都是合式公式，则$(\neg E)$、$(E\vee F)$、$(E\wedge F)$、$(E\rightarrow F)$和$(E\leftrightarrow F)$都是合式公式。

例如，根据基础步骤，就知道**T**、**F**、p和q都是合式公式，其中p和q都是命题变元。从递归步骤的初次应用就知道$(p\vee q)$、$(p\rightarrow \mathbf{F})$、$(\mathbf{F}\rightarrow q)$和$(q\wedge \mathbf{F})$都是合式公式。递归步骤的第二次应用就说明$((p\vee q)\rightarrow(q\wedge \mathbf{F}))$、$(q\vee(p\vee q))$和$((p\rightarrow \mathbf{F})\rightarrow \mathbf{T})$都是合式公式。■

例11　运算与运算数的合式公式　可以递归地定义由变量、数字以及集合$\{+,-,*,/,\uparrow\}$（其中$*$表示乘法且$\uparrow$表示指数）上的运算所组成的合式公式的集合。

基础步骤：若x是数字或变量，则x是合式公式。

递归步骤：若F和G是合式公式，则$(F+G)$、$(F-G)$、$(F*G)$、(F/G)和$(F\uparrow G)$都是合式公式。

例如，根据基础步骤就看出x、y、0和3都是合式公式（因为任何变量和数字都是合式公式）。应用递归步骤一次所生成的合式公式包括$(x+3)$、$(3+y)$、$(x-y)$、$(3-0)$、$(x*3)$、$(3*y)$、$(3/0)$、(x/y)、$(3\uparrow x)$以及$(0\uparrow 3)$等。应用递归步骤两次就说明像$((x+3)+3)$和$(x-(3*y))$这样的公式也是合式公式。[注意$(3/0)$是合式公式，因为在这里只考虑语法。] ■

第 9 章将深入地研究树。树是特殊类型的图，图是由顶点和连接一些顶点对的边所组成的。第 8 章将研究图。在这里将简单地介绍树和图以说明如何来递归地定义树和图。

定义 4 以下这些步骤可以递归地定义根树的集合，其中根树是由一个顶点集合和连接这些顶点的边组成的，顶点集合包含的一个特殊顶点称为树根：

基础步骤：单个顶点 r 是根树。

递归步骤：假设 T_1，T_2，…，T_n 是根树，分别带有树根 r_1，r_2，…，r_n。则如下形成的图也是根树：从树根 r 开始，r 不属于根树 T_1，T_2，…，T_n 中的任何一个，从 r 到顶点 r_1，r_2，…，r_n 中的每个都加入一条边。

图 3-12 解释从基础步骤开始并且应用递归步骤一次和两次所形成的一些根树。注意，递归定义的每次应用都形成了无穷多个根树。

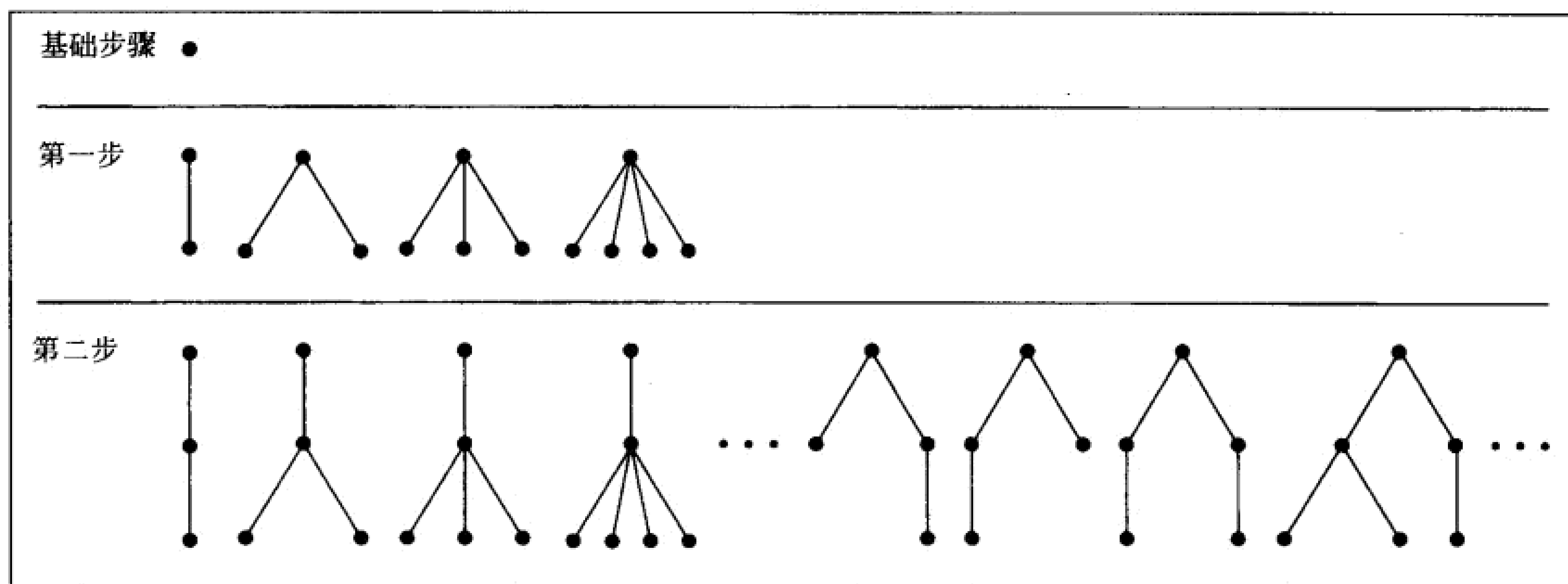


图 3-12 建立根树

根树是特殊情形的二叉树。我们将给出两种类型的二叉树的定义，即满二叉树和扩展二叉树。在每种类型二叉树的定义的递归步骤中，都把两个二叉树组合起来形成一个新的二叉树，指定这两个二叉树中的一个作为左子树，另一个作为右子树。在扩展二叉树中，左子树或右子树可以为空，但是在满二叉树中，这是不允许的。在计算机科学中，二叉树是最重要的结构类型之一。第 9 章将看到二叉树如何用在搜索和排序算法、数据压缩算法以及许多其他应用中。首先定义扩展二叉树。

定义 5 以下这些步骤可以递归地定义扩展二叉树的集合：

基础步骤：空集合是扩展二叉树。

递归步骤：如果 T_1 和 T_2 都是扩展二叉树，则存在一个表示为 $T_1 \cdot T_2$ 的扩展二叉树，它包含树根 r 和当左子树 T_1 和右子树 T_2 都非空时，连接从 r 到这两个子树各自的根的边。

图 3-13 解释如何通过应用递归步骤一次至三次来建立扩展二叉树。

现在说明如何定义满二叉树的集合。注意，这个递归定义与扩展二叉树的递归定义之间的差别完全在于基础步骤。

定义 6 以下这些步骤可以递归地定义满二叉树的集合：

基础步骤：存在一个只含有单个顶点的满二叉树。

递归步骤：如果 T_1 和 T_2 都是满二叉树，则存在一个表示为 $T_1 \cdot T_2$ 的满二叉树，它包含树根 r 和连接从 r 到左子树 T_1 和右子树 T_2 各自的根的边。

图 3-14 解释如何通过应用递归步骤一次和两次来建立满二叉树。

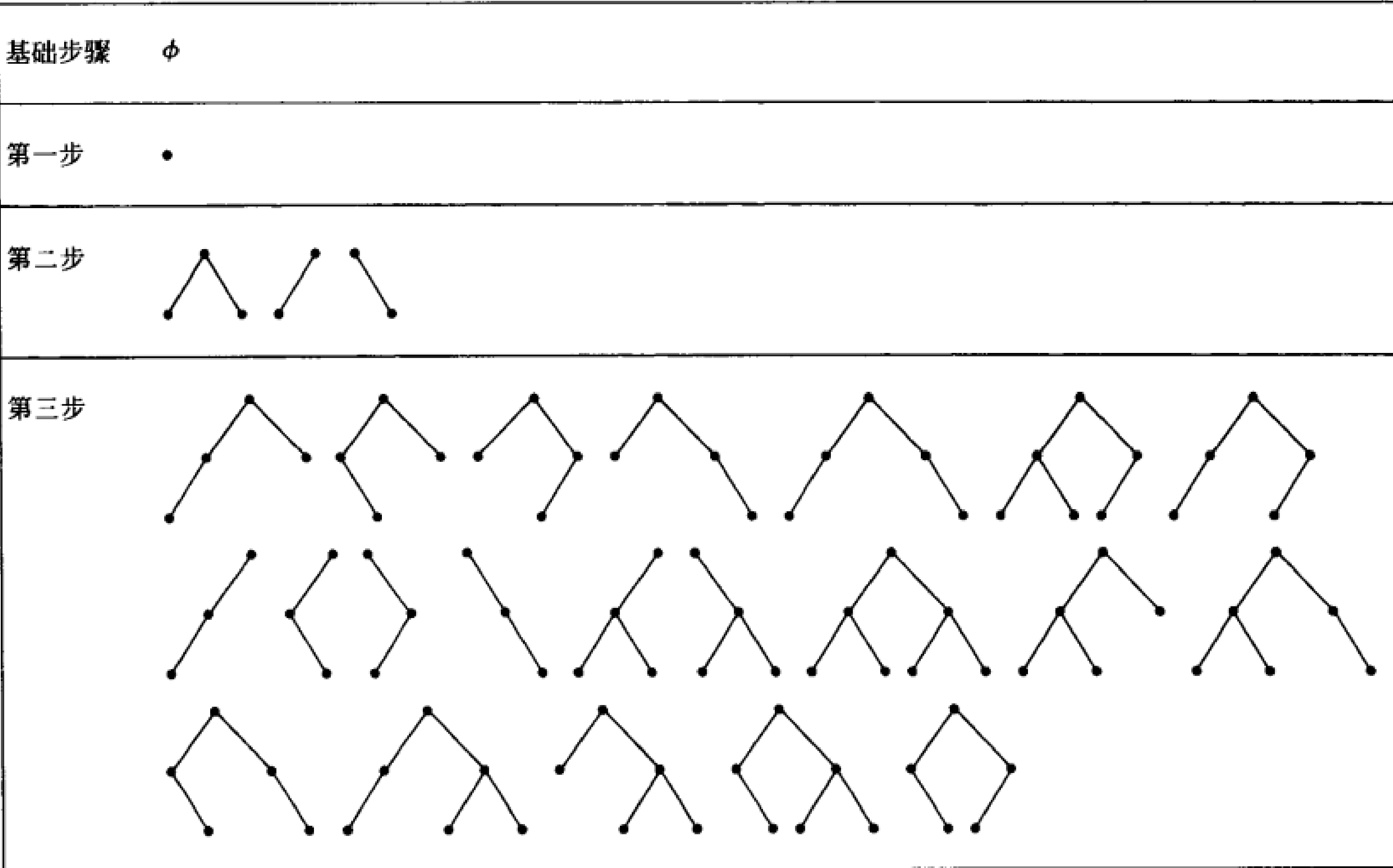


图 3-13 建立扩展二叉树

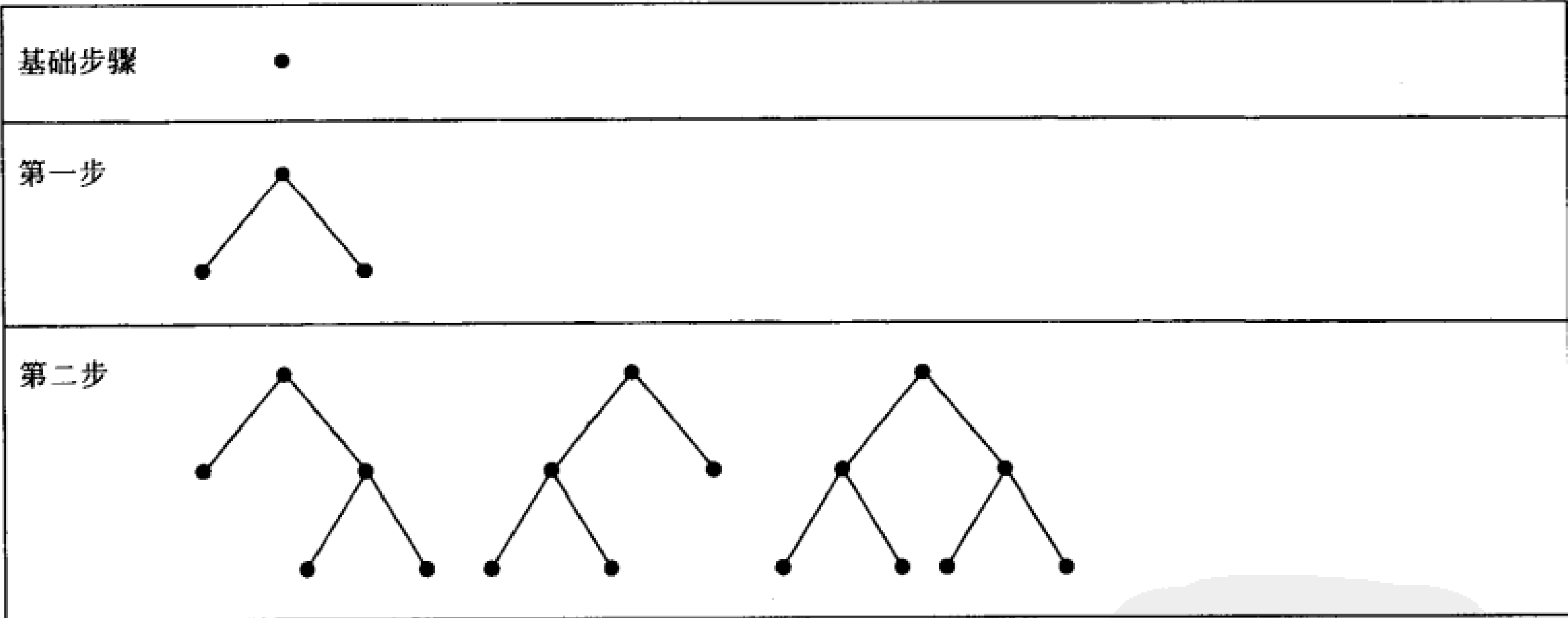


图 3-14 建立满二叉树

3.4.4 结构归纳法

为了证明关于递归定义的集合的结果，一般都使用某种形式的数学归纳法。例 12 解释在递归定义的集合与数学归纳法之间的关系。

例 12 证明例 7 所定义集合 S 是所有为 3 的倍数的正整数的集合。

解 设 A 是被 3 整除的所有正整数的集合。为了证明 $A=S$，必须证明 A 是 S 的子集而且 S 是 A 的子集。为了证明 A 是 S 的子集，必须证明被 3 整除的每个正整数都属于 S。我们将要用数学归纳法来证明它。

设 $P(n)$是命题：$3n$ 属于 S。基础步骤成立，因为根据 S 的递归定义的第一部分，$3\cdot1=3$ 是属于 S 的。为了证明归纳步骤，假定 $P(k)$为真，即 $3k$ 属于 S。因为 $3k$ 属于 S 而且因为 3 属于 S，所以从 S 的递归定义的第二部分得出 $3k+3=3(k+1)$也属于 S。

为了证明 S 是 A 的子集，使用 S 的递归定义。首先，该定义的基础步骤规定 3 属于 S。因为 $3=3\cdot1$，所以所有在这个步骤里属于 S 的元素都被 3 整除。为了完成证明，必须证明所有用该递归定义的第二部分所生成的属于 S 的元素都属于 A。这包括证明每当 x 和 y 都是 S 中的元素并且假定它们都属于 A 时，就有 $x+y$ 属于 A。现在若 x 和 y 都属于 A，则可以得出 $3\mid x$ 和 $3\mid y$。根据 2.4 节的定理 1，得出 $3\mid x+y$，证毕。 ■

例 12 使用正整数集合上的数学归纳法和递归定义来证明关于递归定义的集合的结果。但是，除了直接使用数学归纳法来证明关于递归定义的集合的结果，还有一种更方便的、称为*结构归纳法*的归纳法形式，它不直接使用数学归纳法也可证明关于递归定义的集合的结果。结构归纳法证明包含如下两个部分：

基础步骤：证明对于递归定义的基础步骤所规定的属于该集合的所有元素来说，结果成立。

递归步骤：证明如果对于定义的递归步骤中用来构造新元素的每个元素来说命题为真，则对于这些新的元素来说结果成立。

结构归纳法的有效性来自于非负整数的数学归纳法原理。为了看出这一点，设 $P(n)$为：对于递归定义的递归步骤的不超过 n 次应用所生成的所有元素来说，断言为真。如果能够证明每当 n 是正整数时 $P(n)$就为真，那将证明了数学归纳法原理蕴含着结构归纳法原理。在结构归纳法的基础步骤中，证明 $P(0)$为真。换句话说，证明对于所有在定义的基础步骤中规定为属于集合的元素来说，结果为真。归纳步骤的后果是，如果假设 $P(k)$为真，则得出 $P(k+1)$为真。当用结构归纳法完成一个证明时，就已经证明了 $P(0)$为真并且 $P(k)$蕴含 $P(k+1)$。根据数学归纳法就得出对于所有非负整数 n 来说，$P(n)$为真。这也证明了对于递归定义生成的所有元素来说结果为真，并且证明了结构归纳法是一种有效的证明技术。

使用结构归纳法证明的例子 为了使用结构归纳法来证明关于例 10 所定义的合式表达式集合的结果，需要完成下面的基础步骤和递归步骤。

基础步骤：证明每当 p 是命题变元时，对于 **T**、**F** 和 p 来说结果为真。

递归步骤：证明如果对于复合命题 p 和 q 来说，结果为真，则对于$(\neg p)$、$(p\vee q)$、$(p\wedge q)$、$(p\rightarrow q)$和$(p\leftrightarrow q)$来说，结果都为真。

例 13 解释如何用结构归纳法证明关于合式公式的结果。

例 13 证明例 10 所定义的复合命题的每个合式公式都含有相等个数的左括号和右括号。 ■

证 *基础步骤*：公式 **T**、**F** 和 p 每个都不包含括号，所以显然它们含有相等个数的左括号和右括号。

递归步骤：假设 p 和 q 都是各自含有相等个数的左括号和右括号的合式公式。换句话说，如果 l_p 和 l_q 分别是 p 和 q 中左括号的个数，r_p 和 r_q 分别是 p 和 q 中右括号的个数，则 $l_p=r_p$ 并且 $l_q=r_q$。为了完成归纳步骤，需要证明$(\neg p)$、$(p\vee q)$、$(p\wedge q)$、$(p\rightarrow q)$和$(p\leftrightarrow q)$也各自含有相等个数的左括号和右括号。这些复合命题中第一个的左括号个数等于 l_p+1，其他每个复合命题的左括号个数等于 l_p+l_q+1。同样，这些复合命题中第一个的右括号个数等于 r_p+1，其他每个复合命题的右括号个数等于 r_p+r_q+1。因为$l_p=r_p$并且 $l_q=r_q$，所以这些复合表达式每个都含有相等个数的左括号和右括号。这样就完成了归纳证明。 □

假设 $P(w)$是 $w\in\Sigma^*$ 的字符串集合上的命题函数。为了用结构归纳法来证明对于所有

$w\in\Sigma^*$ 的字符串来说 $P(w)$ 成立，需要同时完成基础步骤和递归步骤。这些步骤是：

基础步骤：证明 $P(\lambda)$ 为真。

递归步骤：假设 $P(w)$ 为真，其中 $w\in\Sigma^*$，证明如果 $x\in\Sigma$，则 $P(wx)$ 也必定为真。

例 14 解释如何在关于字符串的证明中使用结构归纳法。

例 14 用结构归纳法证明 $l(xy)=l(x)+l(y)$，其中 x 和 y 属于 Σ^*，即字母表 Σ 上的字符串的集合。

解 本证明将基于定义 2 所给出对集合 Σ^* 的递归定义和例 9 对字符串长度的定义。设 $P(y)$ 是命题：每当 $x\in\Sigma^*$ 时就有 $l(xy)=l(x)+l(y)$。

基础步骤：为了完成基础步骤，必须证明 $P(\lambda)$ 为真。即必须证明对所有 $x\in\Sigma^*$ 来说，有 $l(x\lambda)=l(x)+l(\lambda)$。因为对每个字符串 x 来说，$l(x\lambda)=l(x)=l(x)+0=l(x)+l(\lambda)$，所以 $P(\lambda)$ 为真。

归纳步骤：为了完成归纳步骤，假定 $P(y)$ 为真，而且证明这个假定蕴含着每当 $a\in\Sigma$ 时，就有 $P(ya)$ 为真。需要证明的是，对每个 $a\in\Sigma$ 来说有 $l(xya)=l(x)+l(ya)$。为了证明这一点，注意到根据 $l(w)$ 的递归定义（在例 9 里给出），有 $l(xya)=l(xy)+1$ 和 $l(ya)=l(y)+1$。而且，根据归纳假设，有 $l(xy)=l(x)+l(y)$。故得出 $l(xya)=l(x)+l(y)+1=l(x)+l(ya)$。 ■

可以用结构归纳法证明关于树或特殊类别的树的结果。例如，为了用结构归纳法证明关于满二叉树的结果，需要证明下面的基础步骤和递归步骤。

基础步骤：证明对于只含有单个顶点的树来说，结果为真。

递归步骤：证明如果对于树 T_1 和 T_2 来说，结果为真，则对于包含树根 r 和以 T_1 作为左子树并且以 T_2 作为右子树的树 $T_1\cdot T_2$ 来说结果为真。

在给出例子说明如何用结构归纳法来证明关于满二叉树的结果之前，需要一些定义。将递归地定义满二叉树 T 的高度 $h(T)$ 和顶点数 $n(T)$。首先定义满二叉树的高度。

定义 7 递归地定义满二叉树 T 的高度 $h(T)$。

基础步骤：只含有树根 r 的满二叉树 T 的高度是 $h(T)=0$。

递归步骤：如果 T_1 和 T_2 都是满二叉树，则满二叉树 $T=T_1\cdot T_2$ 有高度 $h(T)=1+\max(h(T_1),\ h(T_2))$。

如果设 $n(T)$ 表示满二叉树 T 的顶点个数，则注意到 $n(T)$ 满足下面的递归定义：

基础步骤：只含有树根 r 的满二叉树 T 的顶点数 $n(T)$ 是 $n(T)=1$。

递归步骤：如果 T_1 和 T_2 都是满二叉树，则满二叉树 $T=T_1\cdot T_2$ 的顶点数是 $n(T)=1+n(T_1)+n(T_2)$。

现在说明如何可以用结构归纳法来证明一个关于满二叉树的结果。

定理 2 如果 T 是满二叉树，则 $n(T)\leqslant 2^{h(T)+1}-1$。

证 用结构归纳法证明这个不等式。

基础步骤：对于只含有树根 r 的满二叉树来说，结果为真，因为 $n(T)=1$ 并且 $h(T)=0$，所以 $n(T)=1\leqslant 2^{0+1}-1=1$。

归纳步骤：对于归纳假设，假定每当 T_1 和 T_2 都是满二叉树时，$n(T_1)\leqslant 2^{h(T_1)+1}-1$ 并且 $n(T_2)\leqslant 2^{h(T_2)+1}-1$。根据 $n(T)$ 和 $h(T)$ 的递归公式，就有 $n(T)=1+n(T_1)+n(T_2)$ 和 $h(T)=1+\max(h(T_1),\ h(T_2))$。

我们发现

$$\begin{aligned} n(T) &= 1+n(T_1)+n(T_2) && \text{（根据 } n(T) \text{ 的递归公式）} \\ &\leqslant 1+(2^{h(T_1)+1}-1)+(2^{h(T_2)+1}-1) && \text{（根据归纳假设）} \end{aligned}$$

$= 2 \cdot \max(2^{h(T_1)+1}, 2^{h(T_2)+1}) - 1$　　（因为两项之和至多是较大项的 2 倍）

$= 2 \cdot 2^{\max(h(T_1), h(T_2))+1} - 1$

$= 2 \cdot 2^{h(T)} - 1$　　（根据 $h(T)$ 的递归定义）

$= 2^{h(T)+1} - 1$

这样就完成了归纳步骤。□

3.4.5 广义归纳法

可以扩展数学归纳法来证明关于除整数集合以外的其他具有良序性的集合的结果。虽然 7.6 节将讨论这个概念，但是这里给出一个例子来说明这样一种方法的有用性。

作为一个例子，注意到可以定义 $\mathbf{N}\times\mathbf{N}$（非负整数的有序对）上的序，规定如果 $x_1 < x_2$ 或者 $x_1 = x_2$ 且 $y_1 < y_2$，则 (x_1, y_1) 小于等于 (x_2, y_2)。这称为*字典序*。具有这个序的集合 $\mathbf{N}\times\mathbf{N}$ 具有性质：$\mathbf{N}\times\mathbf{N}$ 的每个子集合都有最小元（参见第 7 章补充练习 47）。这意味着可以递归地定义满足 $m\in\mathbf{N}$ 和 $n\in\mathbf{N}$ 的项 $a_{m,n}$，并且用数学归纳法的变种来证明关于这些项的结果，例 15 说明这一点。

例 15　假设对于 $(m, n)\in\mathbf{N}\times\mathbf{N}$ 来说递归地定义 $a_{m,n}$，令 $a_{0,0}=0$ 并且

$$a_{m,n} = \begin{cases} a_{m-1,n} + 1 & \text{若 } n = 0 \text{ 且 } m > 0 \\ a_{m,n-1} + n & \text{若 } n > 0 \end{cases}$$

证明对于所有 $(m, n)\in\mathbf{N}\times\mathbf{N}$ 来说（即对于非负整数的所有有序对来说），$a_{m,n} = m + n(n+1)/2$。

解　用广义的数学归纳法可以证明 $a_{m,n} = m + n(n+1)/2$。基础步骤要求证明：当 $(m, n) = (0, 0)$ 时这个公式是有效的。归纳步骤要求证明：如果对于按照 $\mathbf{N}\times\mathbf{N}$ 上字典序小于 (m, n) 的所有有序对来说，这个公式是有效的，则对于 (m, n) 来说这个公式也是成立的。

基础步骤：设 $(m, n) = (0, 0)$。于是根据 $a_{m,n}$ 递归定义的基础情形就有 $a_{0,0}=0$。另外，当 $m=n=0$ 时，$m+n(n+1)/2 = 0+(0\cdot 1)/2 = 0$。这样就完成了基础步骤。

归纳步骤：假设每当按照 $\mathbf{N}\times\mathbf{N}$ 上字典序 (m', n') 小于 (m, n) 时 $a_{m',n'} = m' + n'(n'+1)/2$。根据递归定义，如果 $n=0$，则 $a_{m,n} = a_{m-1,n}+1$。因为 $(m-1, n)$ 小于 (m, n)，所以归纳假设说 $a_{m-1,n} = m-1+n(n+1)/2$，所以 $a_{m,n} = m-1+n(n+1)/2+1 = m+n(n+1)/2$，给出了所需要的等式。现在假设 $n>0$，所以 $a_{m,n} = a_{m,n-1}+n$。因为 $(m, n-1)$ 小于 (m, n)，所以归纳假设说 $a_{m,n-1} = m+(n-1)n/2$，所以 $a_{m,n} = m+(n-1)n/2+n = m+(n^2-n+2n)/2 = m+n(n+1)/2$。这样就完成了归纳步骤。■

已经说过，在 7.6 节将说明这种证明技术的合理性。

练习

1. 求出 $f(1)$、$f(2)$、$f(3)$ 和 $f(4)$，若 $f(n)$ 递归地定义成：$f(0)=1$，而且对 $n=0, 1, 2, \cdots$ 来说：
   a) $f(n+1)=f(n)+2$　　b) $f(n+1)=3f(n)$
   c) $f(n+1)=2^{f(n)}$　　d) $f(n+1)=f(n)^2+f(n)+1$
2. 求出 $f(1)$、$f(2)$、$f(3)$、$f(4)$ 和 $f(5)$，若 $f(n)$ 递归地定义成：$f(0)=3$，而且对 $n=0, 1, 2, \cdots$ 来说：
   a) $f(n+1)=-2f(n)$　　b) $f(n+1)=3f(n)+7$
   c) $f(n+1)=f(n)^2-2f(n)-2$　　d) $f(n+1)=3^{f(n)/3}$
3. 求出 $f(2)$、$f(3)$、$f(4)$ 和 $f(5)$，若 f 递归地定义成：$f(0)=-1$，$f(1)=2$，而且对 $n=1, 2, \cdots$ 来说：
   a) $f(n+1)=f(n)+3f(n-1)$　　b) $f(n+1)=f(n)^2f(n-1)$
   c) $f(n+1)=3f(n)^2-4f(n-1)^2$　　d) $f(n+1)=f(n-1)/f(n)$

4. 求出 $f(2)$、$f(3)$、$f(4)$和 $f(5)$，若 f 递归地定义成：$f(0)=f(1)=1$，而且对 $n=1$，2，…来说：

a) $f(n+1)=f(n)-f(n-1)$ b) $f(n+1)=f(n)f(n-1)$

c) $f(n+1)=f(n)^2+f(n-1)^3$ d) $f(n+1)=f(n)/f(n-1)$

5. 确定下列这些所谓的定义是否每个都是从非负整数集合到整数集合的函数 f 的有效递归定义。如果 f 是良定义，则求出当 n 是非负整数时 $f(n)$的一个公式并且证明这个公式是有效的。

a) $f(0)=0$，对于 $n\geqslant 1$ 来说 $f(n)=2f(n-2)$

b) $f(0)=1$，对于 $n\geqslant 1$ 来说 $f(n)=f(n-1)-1$

c) $f(0)=2$，$f(1)=3$，对于 $n\geqslant 2$ 来说 $f(n)=f(n-1)-1$

d) $f(0)=1$，$f(1)=2$，对于 $n\geqslant 2$ 来说 $f(n)=2f(n-2)$

e) $f(0)=1$，如果 n 是奇数且 $n\geqslant 1$ 则 $f(n)=3f(n-1)$，如果 n 是偶数且 $n\geqslant 2$ 则 $f(n)=9f(n-2)$

6. 确定下列这些所谓的定义是否每个都是从非负整数集合到整数集合的函数 f 的有效递归定义。如果 f 是良定义，则求出当 n 是非负整数时$f(n)$的一个公式并且证明这个公式是有效的。

a) $f(0)=1$，对于 $n\geqslant 1$ 来说 $f(n)=-f(n-1)$

b) $f(0)=1$，$f(1)=0$，$f(2)=2$，对于 $n\geqslant 3$ 来说$f(n)=2f(n-3)$

c) $f(0)=0$，$f(1)=1$，对于 $n\geqslant 2$ 来说，$f(n)=2f(n+1)$

d) $f(0)=0$，$f(1)=1$，对于 $n\geqslant 1$ 来说，$f(n)=2f(n-1)$

e) $f(0)=2$，如果 n 是奇数且 $n\geqslant 1$ 则 $f(n)=f(n-1)$，如果 $n\geqslant 2$ 则 $f(n)=2f(n-2)$

7. 给出序列$\{a_n\}$的递归定义，$n=1$，2，3，…，若

a) $a_n=6n$ b) $a_n=2n+1$ c) $a_n=10^n$ d) $a_n=5$

8. 给出序列$\{a_n\}$的递归定义，$n=1$，2，3，…，若

a) $a_n=4n-2$ b) $a_n=1+(-1)^n$ c) $a_n=n(n+1)$ d) $a_n=n^2$

9. 设 F 是这样的函数，使得 $F(n)$是前 n 个正整数之和。给出 $F(n)$的递归定义。

10. 给出 $S_m(n)$的递归定义，即整数 m 与非负整数 n 之和。

11. 给出 $P_m(n)$的递归定义，即整数 m 与非负整数 n 之积。

在练习 12～19 里 f_n 是第 n 个斐波那契数。

12. 证明：每当 n 是正整数时，有 $f_1^2+f_2^2+\cdots+f_n^2=f_nf_{n+1}$。

13. 证明：每当 n 是正整数时，有 $f_1+f_3+\cdots+f_{2n-1}=f_{2n}$。

*14. 证明：每当 n 是正整数时，有 $f_{n+1}f_{n-1}-f_n^2=(-1)^n$。

*15. 证明：每当 n 是正整数时，有 $f_0f_1+f_1f_2+\cdots+f_{2n-1}f_{2n}=f_{2n}^2$。

*16. 证明：每当 n 是正整数时，有 $f_0-f_1+f_2-\cdots-f_{2n-1}+f_{2n}=f_{2n-1}-1$。

17. 确定欧几里得算法求出斐波那契数 f_n 和 f_{n+1} 的最大公因子所用的除法次数，其中 n 是非负整数。用数学归纳法验证你的答案。

18. 设

$$\mathbf{A}=\begin{bmatrix}1 & 1\\ 1 & 0\end{bmatrix}$$

证明：每当 n 是正整数时，就有

$$\mathbf{A}^n=\begin{bmatrix}f_{n+1} & f_n\\ f_n & f_{n-1}\end{bmatrix}$$

19. 通过在练习 18 里等式的两边取行列式，证明练习 14 里给出的恒等式。(本题依赖于 2×2 矩阵的行列式概念。)

*20. 给出函数 max 和 min 的递归定义，使得 $\max(a_1, a_2, \cdots, a_n)$和 $\min(a_1, a_2, \cdots, a_n)$分别是 n 个数 a_1，a_2，…，a_n 中的最大值和最小值。

*21. 设 a_1，a_2，…，a_n 和 b_1，b_2，…，b_n 都是实数。用你在练习 20 里给出的递归定义来证明下面的结果。

a) $\max(-a_1, -a_2, \cdots, -a_n) = -\min(a_1, a_2, \cdots, a_n)$

b) $\max(a_1+b_1, a_2+b_2, \cdots, a_n+b_n) \leqslant \max(a_1, a_2, \cdots, a_n) + \max(b_1, b_2, \cdots, b_n)$

c) $\min(a_1+b_1, a_2+b_2, \cdots, a_n+b_n) \geqslant \min(a_1, a_2, \cdots, a_n) + \min(b_1, b_2, \cdots, b_n)$

22. 证明：集合 S 是正整数集合，它定义成：$1 \in S$，而且每当 $s \in S$ 和 $t \in S$ 时就有 $s+t \in S$。
23. 给出是 5 的倍数的正整数集合的递归定义。
24. 给出下述集合的递归定义：

    a)正奇数集合。 b)3 的正整数次幂的集合。 c)整系数多项式的集合。

25. 给出下述集合的递归定义：

    a)正偶数集合。 b)模 3 与 2 同余的正整数的集合。 c)不能被 5 整除的正整数的集合。

26. 设 S 是如下递归地定义的整数有序对的集合：

    基础步骤：$(0, 0) \in S$。

    递归步骤：如果 $(a, b) \in S$，则 $(a+2, b+3) \in S$ 且 $(a+3, b+2) \in S$。

    a)列出递归定义的前 5 次应用所产生的 S 的元素。

    b)对定义的递归步骤的应用次数使用强归纳法来证明当 $(a, b) \in S$ 时 $5 \mid a+b$。

    c)用结构归纳法证明当 $(a, b) \in S$ 时 $5 \mid a+b$。

27. 设 S 是如下递归地定义的整数有序对的集合：

    基础步骤：$(0, 0) \in S$。

    递归步骤：如果 $(a, b) \in S$，则 $(a, b+1) \in S$、$(a+1, b+1) \in S$ 且 $(a+2, b+1) \in S$。

    a)列出递归定义的前 4 次应用所产生的 S 的元素。

    b)对定义的递归步骤的应用次数使用强归纳法来证明当 $(a, b) \in S$ 时 $a \leqslant 2b$。

    c)用结构归纳法证明当 $(a, b) \in S$ 时 $a \leqslant 2b$。

28. 给出下列每个正整数有序对的集合的递归定义。（提示：把集合中的点画在平面上并且寻找包含集合中的点的直线。）

    a) $S = \{(a, b) \mid a \in \mathbf{Z}^+, b \in \mathbf{Z}^+$ 且 $a+b$ 是奇数$\}$

    b) $S = \{(a, b) \mid a \in \mathbf{Z}^+, b \in \mathbf{Z}^+$ 且 $a \mid b\}$

    c) $S = \{(a, b) \mid a \in \mathbf{Z}^+, b \in \mathbf{Z}^+$ 并且 $3 \mid a+b\}$

29. 给出下列每个正整数有序对的集合的递归定义。用结构归纳法证明所找到的递归定义是正确的。（提示：为了找出递归定义，把集合中的点画在平面上并且寻找模式。）

    a) $S = \{(a, b) \mid a \in \mathbf{Z}^+, b \in \mathbf{Z}^+$ 且 $a+b$ 是偶数$\}$

    b) $S = \{(a, b) \mid a \in \mathbf{Z}^+, b \in \mathbf{Z}^+$ 且 a 或 b 是奇数$\}$

    c) $S = \{(a, b) \mid a \in \mathbf{Z}^+, b \in \mathbf{Z}^+$ 且 $a+b$ 是奇数且 $3 \mid b\}$

30. 证明：在位串中，字符串 01 至多比字符串 10 多出现 1 次。
31. 定义由表示集合的变量和 $\{^-, \cup, \cap, -\}$ 里的运算所组成的集合合式公式。
32. a)给出计算位串 s 中 1 的个数的函数 ones(s) 的递归定义。

    b)用结构归纳法证明 ones(st) = ones(s) + ones(t)。

33. a)给出等于十进制数字的非空字符串中最小数字的函数 $m(s)$ 的递归定义。

    b)用结构归纳法证明 $m(st) = \min(m(s), m(t))$。

一个字符串的倒置（反转），是由原字符串里的符号以相反顺序组成的字符串。把字符串 w 的倒置表示成 w^R。

34. 求出下面的位串的倒置。

    a)0101 b)1 1011 c)1000 1001 0111

35. 给出字符串的倒置的递归定义。（提示：首先定义空串的倒置。然后把长度为 $n+1$ 的字符串 w 写成 xy，其中 x 是长度为 n 的字符串，并且利用 x^R 和 y 来表示 w 的倒置。）

*36. 用结构归纳法证明：$(w_1w_2)^R = w_2^R w_1^R$。

37. 给出 w^i 的递归定义，其中 w 是字符串而 i 是非负整数。（在这里 w^i 表示字符串 w 的 i 份复制品的连接。）

*38. 给出回文位串的集合的递归定义。

39. 位串集合 A 定义成

$$\lambda \in A$$
$$\text{若 } x \in A \text{ 则 } 0x1 \in A,$$

其中 λ 是空串。哪些字符串属于 A？

*40. 递归地定义：所包含的0比1多的位串的集合。

41. 用练习37和数学归纳法证明：$l(w^i) = i \cdot l(w)$，其中 w 是位串而 i 是非负整数。

*42. 证明：每当 w 是位串而 i 是非负整数时，就有 $(w^R)^i = (w^i)^R$；即证明一个字符串的倒置的 i 次幂是这个字符串的 i 次幂的倒置。

43. 用结构归纳法证明：$n(T) \geqslant 2h(T)+1$，其中 T 是满二叉树，$n(T)$ 等于 T 的顶点数，$h(T)$ 是 T 的高度。可以递归地定义满二叉树的树叶和内点。

基础步骤：树根 r 是恰有一个顶点 r 的满二叉树的树叶。这个树没有内点。

递归步骤：树 $T = T_1 \cdot T_2$ 的树叶集合是 T_1 的树叶集合与 T_2 的树叶集合之并。T 的内点集合是 T 的树根 r 与 T_1 的内点集合与 T_2 的内点集合之并。

44. 用结构归纳法证明：满二叉树 T 的树叶数 $l(T)$ 比 T 的内点数 $i(T)$ 多1。

45. 仿照例15用广义归纳法证明：如果把 $a_{m,n}$ 递归地定义成 $a_{0,0}=0$ 并且

$$a_{m,n} = \begin{cases} a_{m-1,n}+1 & \text{若 } n=0 \text{ 且 } m>0 \\ a_{m,n-1}+1 & \text{若 } n>0 \end{cases}$$

则对于所有 $(m, n) \in \mathbf{N} \times \mathbf{N}$ 来说 $a_{m,n} = m+n$。

46. 仿照例15用广义归纳法证明：如果把 $a_{m,n}$ 递归地定义成 $a_{1,1}=5$ 并且

$$a_{m,n} = \begin{cases} a_{m-1,n}+2 & \text{若 } n=1 \text{ 且 } m>1 \\ a_{m,n-1}+2 & \text{若 } n>1 \end{cases}$$

则对于所有 $(m, n) \in \mathbf{Z}^+ \times \mathbf{Z}^+$ 来说 $a_{m,n} = 2(m+n)+1$。

*47. 正整数 n 的分拆是把 n 写成正整数之和的方式。例如，$7=3+2+1+1$ 是7的分拆。设 P_m 等于 m 的不同分拆的数目，其中和式里项的顺序无关紧要，并设 $P_{m,n}$ 是用不超过 n 的正整数之和来表示 m 的不同方式数。

a）证明：$P_{m,m} = P_m$。

b）证明：下面的 $P_{m,n}$ 的递归定义是正确的：

$$P_{m,n} = \begin{cases} 1 & \text{若 } m=1 \\ 1 & \text{若 } n=1 \\ P_{m,m} & \text{若 } m<n \\ 1+P_{m,m-1} & \text{若 } m=n>1 \\ P_{m,n-1}+P_{m-n,n} & \text{若 } m>n>1 \end{cases}$$

c）用这个递归定义求出5和6的分拆数。

考虑阿克曼函数的一个变种的下述归纳定义。这个函数是根据德国数学家威尔海姆·阿克曼的名字来命名的，他是大数学家大卫·希尔伯特的学生。在递归函数论以及在涉及集合合并的某些算法的复杂性研究中，阿克曼函数都起到重要的作用。（这个函数有多种不同的变种，都称为阿克曼函数，并且都有类似的性质，尽管它们的值不一定相等。）

$$A(m, n)=\begin{cases}2n & 若 m=0\\ 0 & 若 m\geqslant 1 且 n=0\\ 2 & 若 m\geqslant 1 且 n=1\\ A(m-1, A(m, n-1)) & 若 m\geqslant 1 且 n\geqslant 2\end{cases}$$

练习48～55涉及这种形式的阿克曼函数。

48. 求出阿克曼函数的下列值。

a)$A(1, 0)$　　b)$A(0, 1)$　　c)$A(1, 1)$　　d)$A(2, 2)$

49. 证明：每当 $m\geqslant 1$ 时，就有 $A(m, 2)=4$。

50. 证明：每当 $n\geqslant 1$ 时，就有 $A(1, n)=2^n$。

51. 求出阿克曼函数的下列值。

a)$A(2, 3)$　　* b)$A(3, 3)$

*52. 求出 $A(3, 4)$。

**53. 证明：每当 m 和 n 都是非负整数时，就有 $A(m, n+1)>A(m, n)$。

*54. 证明：每当 m 和 n 都是非负整数时，就有 $A(m+1, n)\geqslant A(m, n)$。

55. 证明：每当 i 和 j 都是非负整数时，就有 $A(i, j)\geqslant j$。

☞56. 用数学归纳法证明：通过规定 $F(0)$ 和从 $F(n)$ 获得 $F(n+1)$ 的规则，所定义的函数 F 是良定义的。

☞57. 用数学归纳法第二原理证明：通过规定 $F(0)$，以及对 $k=0, 1, 2, \cdots, n$，从 $F(k)$ 获得 $F(n+1)$ 的规则，所定义的函数是良定义的。

58. 证明：下述每一个所谓的对正整数集合上的函数的递归定义都不能产生良定义的函数。

a)对 $n\geqslant 1$ 来说 $F(n)=1+F(\lfloor n/2\rfloor)$，而且 $F(1)=1$。

b)对 $n\geqslant 2$ 来说 $F(n)=1+F(n-3)$，而且 $F(1)=2$ 和 $F(2)=3$。

c)对 $n\geqslant 2$ 来说 $F(n)=1+F(n/2)$，而且 $F(1)=1$ 和 $F(2)=2$。

d)若 n 是偶数而且 $n\geqslant 2$ 则 $F(n)=1+F(n/2)$，若 n 是奇数则 $F(n)=1-F(n-1)$，以及 $F(1)=1$。

e)若 n 是偶数而且 $n\geqslant 2$ 则 $F(n)=1+F(n/2)$，若 n 是奇数而且 $n\geqslant 3$ 则 $F(n)=F(3n-1)$，以及 $F(1)=1$。

59. 证明：下述每一个所谓的对正整数集合上的函数的递归定义都不能产生良定义的函数。

a)对 $n\geqslant 1$ 来说 $F(n)=1+F(\lfloor (n+1)/2\rfloor)$，而且 $F(1)=1$。

b)对 $n\geqslant 2$ 来说 $F(n)=1+F(n-2)$，而且 $F(1)=0$。

c)对 $n\geqslant 3$ 来说 $F(n)=1+F(n/3)$，而且 $F(1)=1$，$F(2)=2$，以及 $F(3)=3$。

d)若 n 是偶数而且 $n\geqslant 2$ 则 $F(n)=1+F(n/2)$，若 n 是奇数则 $F(n)=1+F(n-2)$，而且 $F(1)=1$。

e)若 $n\geqslant 2$ 则 $F(n)=1+F(F(n-1))$，而且 $F(1)=2$。

练习60～62处理对数函数的迭代。像通常一样，设 $\log n$ 表示以2为底 n 的对数。函数 $\log^{(k)} n$ 递归地定义成

$$\log^{(k)} n=\begin{cases}n & 若 k=0\\ \log(\log^{(k-1)} n) & 若 \log^{(k-1)} n 有定义且为正数\\ 无定义 & 其他情况\end{cases}$$

迭代对数是函数 $\log^* n$，它在 n 处的值是使得 $\log^{(k)} n\leqslant 1$ 的最小的非负整数 k。

60. 求出下述的每一个值：

a)$\log^{(2)} 16$　　b)$\log^{(3)} 256$　　c)$\log^{(3)} 2^{65\,536}$　　d)$\log^{(4)} 2^{2^{65\,536}}$

61. 对下述的每一个 $\log^* n$ 的值，求出 n 的值：

a)2　b)4　c)8　d)16　e)256　f)65 536　g)$2^{2\,048}$

62. 求出使得 $\log^* n=5$ 的最小整数 n。确定这个数的十进制位数。

练习63～65处理迭代函数的值。假定$f(n)$是从实数集合，或正实数集合，或某些其他的实数集合到实数集合的函数，使得$f(n)$是单调递增的[即当$n<m$时有$f(n)<f(m)$，并且对f的定义域里的所有n来说，$f(n)<n$。]函数$f^{(k)}(n)$递归地定义成

$$f^{(k)}(n)=\begin{cases} n & 若 k=0 \\ f(f^{(k-1)}(n)) & 若 k>0 \end{cases}$$

另外，设c是正实数。*迭代函数*f_c^*是为了把f的自变量缩小到小于或等于c所需要的f的迭代次数，所以$f_c^*(n)$是使得$f^{(k)}(n)\leqslant c$的最小的非负整数k。

63. 设$f(n)=n-a$，其中a是正整数。求出$f^{(k)}(n)$的公式。当n是正整数时，$f_0^*(n)$的值是什么？
64. 设$f(n)=n/2$。求出$f^{(k)}(n)$的公式。当n是正整数时，$f_1^*(n)$的值是什么？
65. 设$f(n)=\sqrt{n}$。求出$f^{(k)}(n)$的公式。当n是正整数时，$f_2^*(n)$的值是什么？

3.5 递归算法

3.5.1 引言

有时可以把带有具体的一组输入的问题的解归约到带更小的一组输入的相同问题的解。例如，求两个正整数a和b的最大公因子的问题，其中$b>a$，就可以归约到求一对更小的整数（即$b\ \mathbf{mod}\ a$和a）的最大公因子的问题，因为$\gcd(b\ \mathbf{mod}\ a,\ a)=\gcd(a,\ b)$。当可以实现这样的归约时，就可以用一系列归约来求出原问题的解，直到把问题归约到解是已知的某个初始情形为止。例如，对求最大公因子来说，归约持续到两个数中较小的一个为零，因为当$a>0$时，$\gcd(a,\ 0)=a$。

我们将看到，连续地把问题归约到带更小的输入的相同问题，这样的算法可用来解决广泛的问题。

定义1 若一个算法通过把问题归约到带更小的输入的相同问题的实例，来解决原来的问题，则这个算法称为*递归的*。

在例1、例2、例4、例5和例6里将描述几个不同的递归算法。第一个例子说明如何构造递归算法从函数的递归定义来求函数的值。

例1 给出计算a^n的递归算法，其中a是非零实数而n是非负整数。

解 可以让递归算法是基于a^n的递归定义。这个定义说对$n>0$来说有$a^{n+1}=a\cdot a^n$，而初始条件是$a^0=1$。为了求出a^n，就连续地用这个递归定义来缩小指数，直到指数是零为止。在算法1里给出这个过程。■

算法1 计算a^n的递归算法

```
procedure power(a：非零实数，n：非负整数)
if n=0 then power(a, n) :=1
else power(a, n) :=a · power(a, n−1)
```

例2 设计一个计算$b^n\ \mathbf{mod}\ m$的递归算法，其中b、n和m是满足$m\geqslant 2$、$n\geqslant 0$且$1\leqslant b<m$的整数。

解 可以基于$b^n\ \mathbf{mod}\ m=(b\cdot(b^{n-1}\ \mathbf{mod}\ m))\ \mathbf{mod}\ m$这个事实来构建递归算法，这个事实来自3.1节练习23和初始条件$b^0\ \mathbf{mod}\ m=1$。在本节末把这个事实留给读者作为练习6。

但是，可以设计出一个更有效的递归算法，用伪码把这个算法描述成算法2。■

算法 2 递归模指数

procedure $mpower(b, n, m$：满足 $m \geq 2$，$n \geq 0$ 且 $1 \leq b < m$ 的整数）

if $n=0$ **then**

  $mpower(b, n, m)=1$

else if n 是偶数 **then**

  $mpower(b, n, m)=mpower(b, n/2, m)^2 \mathbf{mod} m$

else

  $mpower(b, n, m)=(mpower(b, \lfloor n/2 \rfloor, m)^2 \mathbf{mod} m \cdot b \mathbf{mod} m) \mathbf{mod} m$

$\{mpower(b, n, m)=b^n \mathbf{mod} m\}$

可以用强归纳法证明递归算法是正确的，即对于所有可能的输入值来说递归算法都可产生所需要的输出。例 3 通过证明算法 2 是正确的来说明如何做到这一点。

例 3 证明计算模指数的算法 2 是正确的。

Extra Examples

解 对指数 n 用强归纳法。

基础步骤：当 $n=0$ 时 $mpower(b, n, m)=1$。因为每当 b 是整数，m 是整数，满足 $m \geq 2$ 且 $1 \leq b < m$ 时 $b^0 \mathbf{mod}\ m=1$，所以基础步骤就完成了。

归纳步骤：归纳假设是：每当 b 是正整数，m 是整数，满足 $m \geq 2$ 且 $b < m$ 时，对于所有整数 $0 \leq j < k$ 来说，$mpower(b, j, m)=b^j \mathbf{mod}\ m$。

当 k 是偶数时，就有 $mpower(b, k, m)=mpower(b, k/2, m)^2 \mathbf{mod}\ m=(b^{k/2} \mathbf{mod}\ m)^2 \mathbf{mod}\ m=b^k \mathbf{mod}\ m$,其中使用了归纳假设以便把 $mpower(b, k/2, m)$ 换成 $b^{k/2} \mathbf{mod}\ m$。

当 k 是奇数时，就有 $mpower(b, k, m)=((mpower(b, \lfloor k/2 \rfloor, m))^2 \mathbf{mod}\ m \cdot b\ \mathbf{mod}\ m) \mathbf{mod}\ m=((b^{\lfloor k/2 \rfloor} \mathbf{mod}\ m)^2 \mathbf{mod}\ m \cdot b\ \mathbf{mod}\ m) \mathbf{mod}\ m=b^{2\lfloor k/2 \rfloor+1}\ \mathbf{mod}\ m=b^k \mathbf{mod}\ m$，利用了 3.1 节练习 23，因为当 k 是奇数时 $2\lfloor k/2 \rfloor+1=2(k-1)/2+1=k$。这里使用了归纳假设以便把 $mpower(b, \lfloor k/2 \rfloor, m)$ 换成 $b^{\lfloor k/2 \rfloor} \mathbf{mod}\ m$。这样就完成了归纳步骤并且证明算法 2 是正确的。■

下一步给出求最大公因子的递归算法。

例 4 给出求满足 $a < b$ 的两个非负整数 a 和 b 的最大公因子的递归算法。

解 可以基于 $\gcd(a, b)=\gcd(b\ \mathbf{mod}\ a, a)$ 和当 $b>0$ 时 $\gcd(0, b)=b$ 找出算法 3 里的过程。■

算法 3 计算 $\gcd(a, b)$ 的递归算法

procedure $gcd(a, b$：非负整数且 $a < b$）

if $a=0$ **then** $gcd(a, b) := b$

else $gcd(a, b) := gcd(b\ \mathbf{mod}\ a, a)$

现在将要给出 2.1 节所介绍的搜索算法的递归形式。

例 5 把线性搜索算法表达成递归过程。

解 为了在搜索序列 $a_1, a_2, \cdots, a_n$ 里搜索 x，在算法的第 i 步比较 x 与 a_i。若 x 等于 a_i，则 i 是 x 的位置。否则，对 x 的搜索就归约到在少了一个元素的序列（即序列 $a_{i+1}, \cdots, a_n$）里的搜索。现在给出一个递归过程，用伪码把这个过程表示成算法 4。

设 search(i, j, x)是在序列 $a_i, a_{i+1}, \cdots, a_j$ 里搜索 x 的过程。过程的输入包括三元组$(1, n, x)$。若剩余序列的第一项是 x，或者若序列只有一项并且它不是 x，则过程在这一步终止。

若 x 不是第一项而且存在其他的项，则执行同样的过程，但是搜索序列减少一项，它是通过删除搜索序列的第一项而获得的。■

算法4 递归线性搜索算法

```
procedure search(i, j, x)
if a_i = x then
        location := i
else if i = j then
        location := 0
else
        search(i+1, j, x)
```

例6 构造二分搜索算法的递归形式。

解 假定要在序列 a_1，a_2，…，a_n 里求出 x 的位置。为了执行二分搜索，首先比较 x 与中间项 $a_{\lfloor (n+1)/2 \rfloor}$。若 x 等于这一项，则算法将终止。否则，把搜索归约到更小的搜索序列，即若 x 小于原序列的中间项，则归约到序列的前一半，否则归约到后一半。已经把搜索问题的解归约到带长度近似为一半的序列的相同问题的解。二分搜索算法的这种递归形式表达成算法5。■

算法5 递归二分搜索算法

```
procedure binary search(x, i, j)
m := ⌊(i+j)/2⌋
if x = a_m  then
        location := m
else if (x < a_m 且 i < m) then
        binary search(x, i, m-1)
else if (x > a_m 且 j > m) then
        binary search(x, m+1, j)
else location := 0
```

3.5.2 递归与迭代

递归定义把在正整数处的函数值表达成在更小的整数处的函数值。这意味着可以设计递归算法来求出递归地定义的函数在正整数处的值。

例7 算法6中的递归过程当输入是正整数 n 时给出 $n!$ 的值。■

算法6 阶乘的递归过程

```
procedure factorial(n: 正整数)
if n = 1 then
        factorial(n) := 1
else
        factorial(n) := n · factorial(n-1)
```

存在另外一种方式，从阶乘函数的递归定义求它在整数处的值。代替连续地把计算归约到在更小的整数处来求函数的值，可以从在 1 处的函数值开始，连续地应用递归定义来求出在更大的整数处的函数值。这样的过程称为*迭代的*。换句话说，为了用迭代过程求出 $n!$，从 1（即在 1 处的阶乘函数值）开始，连续地乘以每个小于或等于 n 的正整数。这个过程显示在算法 7。

算法 7 阶乘的迭代过程

```
procedure iterative factorial(n: 正整数)
x := 1
for i := 1 to n
        x := i · x
{x 是 n!}
```

在执行了这段代码之后，变量 x 的值是 $n!$。例如，执行这个循环 6 遍之后给出 $6! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 = 720$。

对递归地定义的序列求值的迭代方法，比起使用递归的过程来，常常要求较少量的计算（除非使用专门的递归机器）。用求第 n 个斐波那契数的迭代和递归过程来说明这一点。首先给出递归过程。

算法 8 斐波那契数的递归算法

```
procedure fibonacci(n: 非负整数)
if n=0 then   fibonacci(0) := 0
else if n=1 then   fibonacci(1) := 1
else fibonacci(n) := fibonacci(n-1) + fibonacci(n-2)
```

当使用递归算法求 f_n 时，首先把 f_n 表示成 $f_{n-1} + f_{n-2}$。然后把这两个斐波那契数都换成两个前面的斐波那契数之和。当 f_0 或 f_1 出现时，就直接换成它的值。

注意，在递归的每个阶段，直到获得 f_1 或 f_0 为止，需要求值的斐波那契数的个数都一直翻倍。例如，当使用这个递归算法求出 f_4 时，就必须完成图 3-15 里的树形图所说明的全部计算。这个树包括用 f_4 标记的根，以及从根到用两个斐波那契数 f_3 和 f_2 标记的顶点的分支，它们出现在 f_4 的计算的归约里。每个后续的归约都产生树里的两个分支。当遇到 f_0 和 f_1 时，这种分支结束。读者可以验证一下，这个算法需要 $f_{n+1} - 1$ 次加法来求出 f_n。

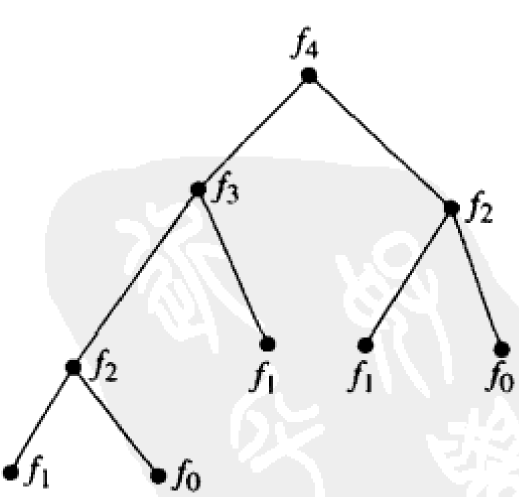


图 3-15 递归地求 f_4 的值

现在考虑用算法 9 中的迭代过程来求出 f_n 所需要的计算量。

算法 9 计算斐波那契数的迭代算法

```
procedure iterative fibonacci(n: 非负整数)
if n=0 then y:=0
else
begin
    x:=0
    y:=1
    for i:=1 to  n-1
    begin
        z:=x+y
        x:=y
        y:=z
    end
end
{y 是第 n 个斐波那契数}
```

这个过程把 x 初始化成 $f_0=0$，把 y 初始化成 $f_1=1$。当经过循环时，把 x 和 y 之和赋给辅助变量 z。然后把 x 赋成 y 的值，而把 y 赋成辅助变量 z 的值。因此，在经过第一次循环之后得出 x 等于 f_1 而 y 等于 $f_0+f_1=f_2$。另外，在经过 $n-1$ 次循环之后 x 等于 f_{n-1} 而且 y 等于 f_n（读者应当验证这个命题）。当 $n>1$ 时，用这个迭代方法求出 f_n 仅仅使用了 $n-1$ 次加法。因此，这个算法比递归算法需要的计算少得多。

已经说明当求递归定义的函数的值时，递归算法可能比迭代算法需要更多的计算量。有时使用递归算法可能更好，即使它比迭代过程更低效，特别是当递归方法容易实现而迭代方法不容易实现时。（另外，或许可以用专门设计来处理递归的机器，它们抵消了使用迭代的好处。）

3.5.3 归并排序

现在描述称为*归并排序*算法的递归排序算法。在概括描述归并排序算法之前，将用一个例子来说明它是如何工作的。

例 8 将用归并排序来排序列表 8，2，4，6，9，7，10，1，5，3。归并排序首先通过不断地把表一分为二来把表分成单个的元素。这个例子的子表的序列表示成图 3-16 上方所示的高度为 4 的平衡二叉树。

排序是通过不断地合并成对的表来完成的。在第一阶段里，把成对的单个元素合并成按升序排列的长度为二的表。然后对成对的表进行相继的合并，直到整个表都排成了升序为止。把这些合并成按升序排列的表的序列表示成图 3-16 下方所示的高度为 4 的平衡二叉树（注意这个树是“上下颠倒地”显示的）。 ■

在一般情况下，归并排序是这样进行的：反复地把表分成长度相等的两个子表（或者其中一个子表比另一个子表多一个元素），直到每个子表包含一个元素为止。这些子表的序列可以表示成平衡的二叉树。这个过程这样继续进行：不断地合并成对的子表，其中的两个表都是按升序排

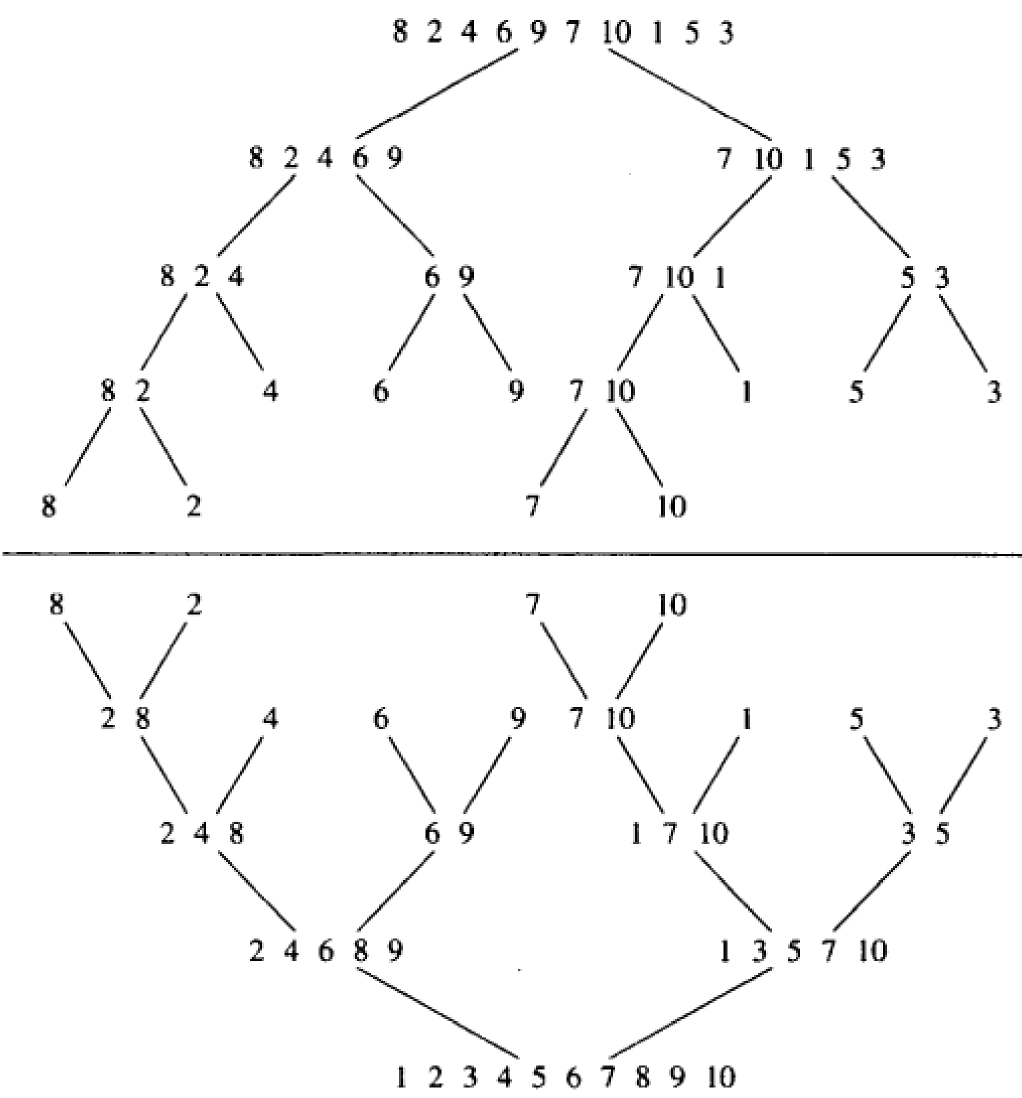


图 3-16 对 8，2，4，6，9，7，10，1，5，3 的归并排序

列的，把它们合并成元素都是按升序排列的较大的表，直到原来的表排成升序为止。这些合并的子表的序列可以表示成平衡的二叉树。

也可以递归地描述归并排序。为了做归并排序，把表分成大小相等或近似相等的两个子表，用归并排序算法排序每个子表，然后合并这两个子表。算法 10 给出归并排序的递归形式。这个算法使用子过程 *merge*，算法 11 描述 *merge*。

> **算法 10** 递归归并排序
>
> **procedure** *mergesort*($L=a_1, \cdots, a_n$)
>
> **if** $n>1$ **then**
>
> $\quad m := \lfloor n/2 \rfloor$
>
> $\quad L_1 := a_1, a_2, \cdots, a_m$
>
> $\quad L_2 := a_{m+1}, a_{m+2}, \cdots, a_n$
>
> $\quad L := merge(mergesort(L_1), mergesort(L_2))$
>
> {现在 L 中的元素以非降顺序排列}

为了实现归并排序，需要把两个有序表合并成更大有序表的有效算法。现在将描述这样的过程。

例 9 将描述如何合并两个表 2，3，5，6 和 1，4。表 3-3 说明所使用的步骤。

表 3-3 合并已排序的表 2，3，5，6 和 1，4

第一个表	第二个表	合并的表	比较
2 3 5 6	1 4		1<2
2 3 5 6	4	1	2<4
3 5 6	4	1 2	3<4
5 6	4	1 2 3	4<5
5 6		1 2 3 4	
		1 2 3 4 5 6	

首先，比较两个表里的最小元素，它们分别是 2 和 1。因为 1 较小，所以把它放在合并的表的开头并且从第二个表删除它。在这个阶段上，第一个表是 2，3，5，6，第二个表是 4，组合而成的表是 1。

其次，比较 2 和 4，它们是两个表里的最小元素。因为 2 较小，所以添加它到组合的表并且从第一个表删除它。在这个阶段上，第一个表是 3，5，6，第二个表是 4，组合而成的表是 1，2。

继续比较 3 和 4，它们是各自表里的最小元素。因为 3 是这两个元素中较小的，所以添加它到组合的表并且从第一个表删除它。在这个阶段上，第一个表是 5，6，第二个表是 4，组合而成的表是 1，2，3。

然后比较 5 和 4，它们是两个表里的最小元素。因为 4 是这两个元素中较小的，所以添加它到组合的表并且从第二个表删除它。在这个阶段上，第一个表是 5，6，第二个表是空的，组合而成的表是 1，2，3，4。

最后，因为第二个表是空的，所以第一个表的所有元素可以附加到组合表的后面，保持它们在第一个表里的出现顺序。这样就产生出有序表 1，2，3，4，5，6。 ■

现在将考虑合并两个有序表 L_1 和 L_2 成一个有序表 L 的一般问题。可以使用下面的过程。从空表 L 开始。比较两个表的最小元素。把这两个元素中较小的放到 L 后面，并且从它所在的表删除它。下一步，若 L_1 和 L_2 有一个是空的，则附加另一个(非空)表到 L，这样就完成了合并。若 L_1 和 L_2 都非空，则重复这个过程。算法 11 给出这个过程的伪代码描述。

```
算法 11  归并两个表
procedure merge(L1, L2: 表)
L := 空表
while L1 和 L2 都非空
begin
    从L1 和 L2 的第一个元素中较小的元素所在的表中删除这个元素并且
      把这个元素放到 L 的左端
    if  删除这个元素导致一个表为空 then 从另一个表中删除所有元素并且把这些元素附加
        到 L 的后面
end{L 是元素按照递增顺序排列的归并的表}
```

在对归并排序的分析中，将需要估计合并两个有序表 L_1 和 L_2 所用的比较次数。对于算法 11 来说，可以容易地得出这样的估计。每次比较 L_1 的一个元素与 L_2 的一个元素，把一个附加元素添加到合并的表 L 中。不过，当 L_1 或 L_2 为空时，就不需要更多的比较了。因此，当执行 $m+n-2$ 次比较时，其中 m 和 n 分别是 L_1 和 L_2 中的元素个数，算法 11 效率最低，L_1 和 L_2 每

个只剩下一个元素。下一次比较将是最后一次，因为这次比较使得这两个表之一为空。因此，算法 11 使用不超过 $m+n-1$ 次比较。下面的引理总结了这个估计。

引理 1 使用不超过 $m+n-1$ 次比较，可以把带 m 个元素和 n 个元素的两个有序表合并成一个有序表。

有时使用远远少于 $m+n-1$ 次比较就可以合并两个长度为 m 和 n 的有序表。例如，当 $m=1$ 时，可以用二叉搜索过程来把第一个表里的这一个元素放进第二个表。这只需要$\lceil \log n \rceil$次比较，对 $m=1$ 来说，$\lceil \log n \rceil$比 $m+n-1=n$ 小得多。在另一方面，对 m 和 n 的某些值来说，引理 1 给出了最好可能的界限。即存在着带有 m 个和 n 个元素的表，比较次数少于 $m+n-1$ 次是无法合并它们的。(见本节末尾的练习 35。)

现在可以分析归并排序的复杂性了。与研究一般性问题不一样的是，将假定表中的元素个数 n 是 2 的幂，比方说 2^m。这样将使得分析不是太复杂，但是当实际情况不是这样时，还可以做各种修改，这些修改将产生同样的估计。

在分解过程的第一阶段上，把表分解成两个子表，每个子表都有 2^{m-1}个元素，位于分解所生成的树的 1 层上。这个过程继续下去，把两个带 2^{m-1}个元素的子表分解成四个在 2 层上各有 2^{m-2}个元素的子表，并依此类推。在一般情况下，在 $k-1$ 层上有 2^{k-1}个表，每个表有 2^{m-k+1}个元素。在 $k-1$ 层上的这些表分解成在 k 层上的 2^k 个表，每个表有 2^{m-k}个元素。在这个过程的最后，有 2^m 个表，每个表有一个元素，在 m 层上。

可以这样来开始合并：把 2^m 个含一个元素的表成对地组合成 2^{m-1}个表，都在 $m-1$ 层上，各有两个元素。为了这样做，把 2^{m-1}对含一个元素的表合并。每一对表的合并恰好需要两次比较。

这个过程继续下去，使得在 k 层上($k=m, m-1, m-2, \cdots, 3, 2, 1$)，$2^k$ 个各有 2^{m-k}个元素的表合并成 2^{k-1}个表，各有 2^{m-k+1}个元素，都在 $k-1$ 层上。为了这样做，需要总共 2^{k-1}次合并两个表，每个表有 2^{m-k}个元素。但是，根据引理 1，这些合并每个都可以用至多 $2^{m-k}+2^{m-k}-1=2^{m-k+1}-1$ 次比较来完成。因此，从 k 层进行到 $k-1$ 层，可以用至多 $2^{k-1}(2^{m-k+1}-1)$ 次比较来完成。对所有这些估计求和就证明了归并排序所需要的比较次数至多是

$$\sum_{k=1}^{m} 2^{k-1}(2^{m-k+1}-1) = \sum_{k=1}^{m} 2^m - \sum_{k=1}^{m} 2^{k-1} = m2^m - (2^m - 1) = n\log n - n + 1$$

因为 $m=\log n$ 和 $n=2^m$。(这样求 $\sum_{k=1}^{m} 2^m$ 的值:注意它是 m 个相同的项之和,每个都等于 2^m。这样求 $\sum_{k=1}^{m} 2^{k-1}$ 的值:用 3.2 节定理 1 几何级数各项求和的公式。)

这样的分析说明,归并排序达到了排序算法所需比较次数的最好可能的大 O 估计,如下面的定理所述。

定理 1 对一个带 n 个元素的表进行归并排序所需要的比较次数是 $O(n\log n)$。

在练习里描述另一个有效的算法——快速排序。

练习

1. 给出每当 n 是正整数而 x 是整数时，计算 nx 的递归算法。
2. 给出求前 n 个正整数之和的递归算法。
3. 给出求前 n 个正奇数之和的递归算法。
4. 给出求一组有限个整数的最大值的递归算法。
5. 给出求一组有限个整数的最小值的递归算法。

6. 设计一个递归算法，每当 n、x 和 m 都是正整数时，基于事实 $x^n \textbf{ mod } m=(x^{n-1} \textbf{ mod } m \cdot x \textbf{ mod } m) \textbf{ mod } m$ 求出 $x^n \textbf{ mod } m$。
7. 给出每当 n 和 m 都是正整数时，求 $n! \textbf{ mod } m$ 的递归算法。
8. 给出求出整数列表中的众数的递归算法。(众数是列表中出现的频繁程度至少与其他每个元素一样的元素。)
9. 设计一个递归算法，它计算假如 $\gcd(a, b)=\gcd(a, b-a)$ 时，满足 $a<b$ 的两个非负整数 a 和 b 的最大公因子。
10. 证明：练习2找到的求前 n 个正整数之和的递归算法是正确的。
11. 设计把两个非负整数 x 和 y 相乘的递归算法，基于这样的事实：当 y 是偶数时 $xy=2(x\cdot(y/2))$，当 y 是奇数时 $xy=2(x\cdot\lfloor y/2\rfloor)+x$，以及初始条件：当 $y=0$ 时 $xy=0$。
12. 证明：练习11设计的算法是正确的。
13. 证明：练习1找到的递归算法是正确的。
14. 证明：练习4找到的递归算法是正确的。
15. 证明：算法1是正确的。
16. 设计求 a^{2^n} 的递归算法，其中 a 是实数而 n 是正整数。[提示：利用等式 $a^{2^{n+1}}=(a^{2^n})^2$。]
17. 对于求 a^{2^n} 的值，练习16里的算法所用的乘法次数与算法1所用的乘法次数相比较的结果如何?
*18. 用练习16里的算法，设计当 n 是非负整数时求 a^n 的值的算法。(提示：利用 n 的二进制展开式。)
*19. 对于求 a^n 的值，练习18里的算法所用的乘法次数与算法1所用的乘法次数相比较的结果如何?
20. 为了求出斐波那契数 f_7，在算法8和算法9里给出的递归算法和迭代算法，各自分别使用多少次加法?
21. 设计求一个序列的第 n 项的递归算法，该序列定义成：$a_0=1$，$a_1=2$，而且对 $n=2, 3, 4, \cdots$ 来说有 $a_n=a_{n-1}\cdot a_{n-2}$。
22. 设计求练习21里定义的序列的第 n 项的迭代算法。
23. 求练习21里的序列的递归算法与迭代算法，哪个算法更有效?
24. 设计求一个序列的第 n 项的递归算法，该序列定义成：$a_0=1$，$a_1=2$，$a_2=3$，而且对 $n=3, 4, 5, \cdots$ 来说有 $a_n=a_{n-1}+a_{n-2}+a_{n-3}$。
25. 设计求练习24里定义的序列的第 n 项的迭代算法。
26. 求练习24里的序列的递归算法与迭代算法，哪个算法更有效?
27. 给出求一个序列的第 n 项的递归算法和迭代算法，该序列定义成：$a_0=1$，$a_1=3$，$a_2=5$，而且 $a_n=a_{n-1}\cdot a_{n-2}^2\cdot a_{n-3}^3$。哪个算法更有效?
28. 根据3.4节练习47里给出的递归定义，给出求正整数的划分数的递归算法。
29. 给出求字符串的倒置的递归算法。(见3.4节练习34前面的说明里对位串的倒置的定义。)
30. 给出当 w 是位串时，求字符串 w^i(即 w 的 i 个复制品的连接)的递归算法。
31. 给出递归算法来计算阿克曼函数的值。(提示：见3.4节练习48前面的说明。)
32. 用归并排序来排序4，3，2，5，1，8，7，6，说明算法所用的所有步骤。
33. 用归并排序来排序 b，d，a，f，g，h，z，p，o，k，说明算法所用的所有步骤。
34. 为了用算法11来合并下面的成对的表，需要多少次比较?
    a)1，3，5，7，9；2，4，6，8，10
    b)1，2，3，4，5；6，7，8，9，10
    c)1，5，6，7，8；2，3，4，9，10
35. 证明：存在着带有 m 个和 n 个元素的表，使得它们不能用算法11以少于 $m+n-1$ 次的比较来合并成一个有序表。
*36. 当两个升序的表里的元素个数如下时，把它们合并成一个升序的表，所需要的最少比较次数是什么?
    a)1，4　　b)2，4　　c)3，4　　d)4，4

*37. 证明：归并排序算法是正确的。

快速排序是一个有效算法。为了排序 a_1，a_2，…，a_n，这个算法首先挑出第一个元素 a_1 并且构造两个子表，第一个子表包含小于 a_1 的元素，是按照元素出现的顺序排列的。第二个子表包含大于 a_1 的元素，是按照元素出现的顺序排列的。然后把 a_1 放在第一个子表的后面。对每个子表递归地重复这个过程，直到所有子表都只包含一个项为止。n 个项的有序表是这样获得的：按照只含有一个项的子表出现的顺序来组合它们。

38. 用快速排序来排序 3，5，7，8，1，9，2，4，6。

39. 设 a_1，a_2，…，a_n 是 n 个不同实数的表。从这个表构造两个子表，第一个子表包含小于 a_1 的元素而第二个子表包含大于 a_1 的元素，那么需要多少次比较？

40. 用伪代码描述快速排序算法。

41. 用快速排序算法来排序四个元素的表，需要的最大比较次数是什么？

42. 用快速排序算法来排序四个元素的表，需要的最小比较次数是什么？

43. 就所用的比较次数而言，确定快速排序算法的最坏情形复杂性。

*44. 证明：对 $n>4$ 来说，$\log n!$ 大于$(n\log n)/4$。(提示：从不等式 $n! > n(n-1)(n-2)\cdots\lceil n/2\rceil$入手。)

3.6 程序正确性

3.6.1 引言

假定设计了一个解决问题的算法，而且编写了实现它的程序。如何才能保证这个程序总是产生正确的答案？在消除了所有的错误使得语法正确之后，可以用简单的输入来测试这个程序。若对任何简单输入来说产生了不正确的结果，则它是不正确的。但是即使对所有的简单输入来说这个程序都给出了正确的答案，它也不一定总是产生正确的答案(除非已经测试了所有可能的输入)。需要一个说明这个程序总是给出正确答案的证明。

程序验证(即程序正确性的证明)使用在本章里描述的推理规则和证明技术，包括数学归纳法。因为不正确的程序可能导致灾难性的后果，所以已经构造了大量的方法来对程序进行验证。在使程序验证自动化以便可以用计算机来完成方面已经作出了大量努力，但取得的进展是非常有限的。事实上，一些数学家和计算机理论家争论的使复杂程序的正确性证明机械化永远是不现实的。

本节将介绍用来证明程序为正确的一些概念和方法。不过，在本书里将不逐步展开讨论程序验证的完整方法。本节将把逻辑规则、证明技术以及算法的概念联系在一起，对程序验证领域给予粗略的介绍。

3.6.2 程序验证

若对每个可能的输入来说一个程序都产生正确的输出，则说这个程序是*正确的*。一个程序的正确证明包括两个部分。第一部分证明：若程序终止，则获得正确的答案。证明的这一部分证明了程序的*部分正确性*。证明的第二部分证明：程序总是终止。

为了规定程序产生正确的输出是什么意思，使用两个命题。第一个是*初始断言*，它给出输入值必须具有的性质。第二个是*终结断言*，它给出假如程序做了要求它做的事情，则程序的输出应当具有的性质。当验证一个程序时，必须提供适当的初始断言和终结断言。

定义 1 若每当对一个程序或程序段 S 的输入值来说初始断言 p 为真时，就有对 S 的输出值来说终结断言 q 为真，则说 S 是*相对于 p 和 q 部分正确的*。记号 $p\{S\}q$ 说明程序或程序段 S 是相对于初始断言 p 和终结断言 q 部分正确的。

注：记号 $p\{S\}q$ 称为霍尔三元组，因为托尼·霍尔[⊖]引入了部分正确性的概念。

注意，部分正确性的概念与程序是否终止是无关的；它仅仅关注若程序终止，则程序是否做了期待它做的事情。

可以用一个简单的例子说明初始断言和终结断言的概念。

例 1 证明程序段

$y:=2$
$z:=x+y$

是相对于初始断言 p：$x=1$ 和终结断言 q：$z=3$ 部分正确的。

解 假定 p 为真，所以在程序开始时 $x=1$。则把 y 赋值成 2，而把 z 赋值成 x 和 y 值之和，即 3。因此，S 是相对于初始断言 p 和终结断言 q 部分正确的。因此，$p\{S\}q$ 为真。 ■

3.6.3 推理规则

一条有用的推理规则通过把一个程序分成一系列子程序，然后证明每个子程序为正确的来证明这个程序为正确的。

假定把程序 S 分成子程序 S_1 和 S_2。写 $S=S_1$；S_2 来表示 S 是由 S_1 后接 S_2 来组成的。假定已经证明了 S_1 相对于初始断言 p 和终结断言 q 的正确性，以及 S_2 相对于初始断言 q 和终结断言 r 的正确性。由此得出了若 p 为真而且 S_1 执行且终止则 q 为真；若 q 为真而且 S_2 执行且终止则 r 为真。因此，若 p 为真而且 $S=S_1$；S_2 执行且终止则 r 为真。这条推理规则称为**合成规则**，它可以叙述成

$$\begin{array}{l} p\{S_1\}q \\ q\{S_2\}r \\ \hline \therefore\ p\{S_1;\ S_2\}r \end{array}$$

在本节后面将使用这条推理规则。

下一步，将给出含有条件语句和循环的程序段的推理规则。因为可以把程序分成程序段，以便进行正确性证明，所以这样就能够验证许多不同的程序。

3.6.4 条件语句

首先将要给出条件语句的推理规则。假定一个程序段形如

```
If condition then
    S
```

其中 S 是一个语句块。若 *condition* 为真，则 S 执行，而当 *condition* 为假时，则 S 不执行。为了验证这个程序段相对于初始断言 p 和终结断言 q 来说是正确的，必须做两件事情。首先，必须证明：当 p 为真而且 *condition* 也为真时，在 S 终止之后 q 为真。其次，必须证明当 p 为真而且

⊖ C·安东尼·R·霍尔(生于 1934 年)，托尼·霍尔(Tony Hoare)目前是英格兰牛津大学计算机科学教授以及皇家协会成员。霍尔对编程语言的理论和编程方法学作出了许多重要贡献。他是第一个基于如何证明程序相对于它们的规定来说为正确来定义编程语言的人。霍尔也是快速排序的发明者，这是最普遍地使用和研究的排序算法之一(见 3.5 节的练习)。霍尔是著名的计算机科学技术层面和社会层面的作家。

condition 为假时，q 为真（因为在这种情形里 S 不执行）。

这导致下面的推理规则：

$(p \wedge condition)\{S\}q$
$(p \wedge \neg condition) \rightarrow q$

$\therefore p\{\textbf{if } condition \textbf{ then } S\}q$

例 2 说明如何使用这条推理规则。

例 2 验证程序段

```
if x>y then
    y:=x
```

是相对于初始断言 **T** 和终结断言 $y \geq x$ 为正确的。

解 当初始断言为真而 $x>y$ 时，则完成赋值语句 $y:=x$。因此，在这种情形里，断言 $y \geq x$ 的终结断言为真。另外，当初始断言为真而 $x>y$ 为假、因而 $x \leq y$ 时，终结断言再次为真。因此，使用这种类型的程序段的推理规则，这个程序是相对于给定的初始断言和终结断言为正确的。■

同理，考虑含有语句形如

```
if condition then
    S1
else
    S2
```

的程序。若 *condition* 为真，则执行 S_1；若 *condition* 为假，则执行 S_2。为了验证这个程序段是相对于初始断言 p 和终结断言 q 为正确的，必须做两件事情。首先，必须证明：当 p 为真而且 *condition* 为真时，在 S_1 终止之后 q 为真。其次，必须证明：当 p 为真而且 *condition* 为假时，在 S_2 终止之后 q 为真。这导致下面的推理规则：

$(p \wedge condition)\{S_1\}q$
$(p \wedge \neg condition)\{S_2\}q$

$\therefore p\{\textbf{if } condition \textbf{ then } S_1 \textbf{ else } S_2\}q$

下面的例子说明如何使用这条推理规则。

例 3 验证程序段

```
if x<0 then
    abs:=-x
else
    abs:=x
```

是相对于初始断言 **T** 和终结断言 $abs=|x|$ 为正确的。

解 必须证明两件事情。首先必须证明：若初始断言为真而 $x<0$ 则 $abs=|x|$。这是正确的，因为当 $x<0$ 时赋值语句 $abs:=-x$ 让 $abs=-x$，根据定义当 $x<0$ 时它是 $|x|$。其次必须证明：若初始断言为真而 $x<0$ 为假（所以 $x \geq 0$）则 $abs=|x|$。这是正确的，因为在这种情形

里，程序使用赋值语句 $abs := x$，而根据定义，当 $x \geqslant 0$ 时 x 是 $|x|$，所以 $abs = |x|$。因此，利用对于这种类型的程序段的推理规则，这个程序是相对于给定的初始断言和终结断言为正确的。■

3.6.5 循环不变量



下一步将要描述 **while** 循环的正确性证明。为了逐步阐述一个

```
while condition
    S
```

这种类型的程序段的推理规则，要注意，S 反复执行直到 *condition* 变假为止。必须选择一个每次执行 S 时都保持为真的断言。这样的断言称为*循环不变量*。换句话说，若 $(p \wedge condition)\{S\}p$ 为真，则 p 是循环不变量。

假定 p 是循环不变量。所以若在执行这个程序段之前 p 为真，则在程序终止后 p 和 $\neg condition$ 都为真，假如程序真的终止。这个推理规则是

$$\frac{(p \wedge condition)\{S\}p}{\therefore p\{\textbf{while}\ condition\ S\}(\neg condition \wedge p)}$$

例 4 需要一个循环不变量来验证当 n 是正整数时，如下程序段以 $factorial = n!$ 终止。

```
i := 1
factorial := 1
while i < n
begin
    i := i + 1
    factorial := factorial · i
end
```



设 p 是命题："$factorial = i!$ 并且 $i \leqslant n$"。首先证明 p 是循环不变量。假设在执行一遍 **while** 循环的开头时，p 为真而且 **while** 循环的条件成立；换句话说，$factorial = i!$ 而且 $i < n$。i 和 $factorial$ 的新值 i_{new} 和 $factorial_{new}$ 现在是 $i_{new} = i + 1$ 和 $factorial_{new} = factorial_{new} \cdot (i+1) = (i+1)! = i_{new}!$。由于 $i < n$，所以也有 $i_{new} = i + 1 \leqslant n$。因此在循环执行的结尾 p 为真。这就证明了 p 是循环不变量。

现在考虑上述程序段。在正好要进入循环之前，$i = 1 \leqslant n$ 和 $factorial = 1 = 1! = i!$ 都为真，所以 p 为真。由于 p 是循环不变量，所以刚刚介绍过的推理规则就蕴含着如果 **while** 循环终止，那么循环终止时 p 为真而且 $i < n$ 为假。在这样的情况下，最终 $factorial = i!$ 和 $i \leqslant n$ 都为真，但 $i < n$ 为假；换句话说，$i = n$ 而且 $factorial = i! = n!$，这正是想要的结果。

最后还需要验证 **while** 循环确实终止。在程序开头把 i 赋值成 1，所以在 $n-1$ 次执行循环后，i 的新的值是 n，循环在这时就终止了。■

下面将给出最后一个例子来说明如何用各种推理规则去验证较长的程序的正确性。

例 5 将要简述如何验证计算两个整数之积的程序 S 的正确性。

```
procedure multiple(m, n: 整数)
     { if n<0 then a := -n
  S1 { else a := n

  S2 { k := 0
     { x := 0

     { while k<a
     { begin
  S3 {      x := x+m
     {      k := k+1
     { end

  S4 { if n<0 then product := -x
     { else product := x
```

目标是证明在执行 S 之后 *product* 有值 mn。通过把 S 分成 $S=S_1;\ S_2;\ S_3;\ S_4$，如 S 的程序清单所示那样，就可以完成正确性证明。可以用合成规则来建立正确性证明。细节将留给读者作为练习。

设 p 是初始断言：m 和 n 都是整数。则可以证明当 q 是命题 $p\wedge(a=|n|)$时，$p\{S_1\}q$ 为真。下一步，设 r 是命题 $q\wedge(k=0)\wedge(x=0)$。容易验证 $q\{S_2\}r$ 为真。可以证明"$x=mk$ 而且 $k\leqslant a$"是 S_3 里的循环不变量。另外，容易看出，在 a 次循环之后循环终止而且 $k=a$，所以在这时 $x=ma$。因为 r 蕴含着 $x=m\cdot 0$ 和 $0\leqslant a$，所以在进入循环之前循环不变量为真。因为循环终止而且 $k=a$，由此得出 $r\{S_3\}s$ 为真，其中 s 是命题"$x=ma$ 而且 $a=|n|$"。最后，可以证明 S_4 是相对于初始断言 s 和终结断言 t 为正确的，其中 t 是命题"$product=mn$"。

把所有这些结果放到一起来考虑，因为 $p\{S_1\}q$、$q\{S_2\}r$、$r\{S_3\}s$ 和 $s\{S_4\}t$ 都为真，所以从合成规则得出 $p\{S\}t$ 为真。另外，因为所有四个程序段都终止，所以 S 终止。这样就验证了这个程序的正确性。 ■

练习

1. 证明程序段

```
y := 1
z := x+y
```

是相对于初始断言 $x=0$ 和终结断言 $z=1$ 为正确的。

2. 验证程序段

```
if x<0 then x := 0
```

是相对于初始断言 T 和终结断言 $x\geqslant 0$ 为正确的。

3. 验证程序段

```
x := 2
z := x+y
if y>0 then
    z := z+1
else
    z := 0
```

是相对于初始断言 $y=3$ 和终结断言 $z=6$ 为正确的。

4. 验证程序段

```
if x<y then
    min := x
else
    min := y
```

是相对于初始断言 T 和终结断言$(x\leqslant y\wedge min=x)\vee(x>y\wedge min=y)$为正确的。

*5. 设计一条推理规则来验证形如

```
if condition 1 then
    S1
else if condition 2 then
    S2
    ⋮
else
    Sn
```

的语句的部分正确性，其中 S_1，S_2，…，S_n 都是语句块。

6. 使用在练习 5 里讨论的推理规则去验证程序：

```
if x<0 then
    y := -2|x|/x
else if x>0 then
    y := 2|x|/x
else if x=0 then
    y := 2
```

是相对于初始断言 T 和终结断言 $y=2$ 为正确的。

7. 用循环不变量证明下述计算实数 x 的 n 次方幂的程序是正确的，其中 n 是正整数。

```
power := 1
i := 1
while i≤n
begin
    power := power * x
    i := i+1
end
```

*8. 证明在 3.5 节里给出的求 f_n 的迭代程序是正确的。

9. 给出在例 5 里给出的正确性证明的所有细节。

10. 假定蕴含式 $p_0\rightarrow p_1$ 和程序断言 $p_1\{S\}q$ 都为真。证明 $p_0\{S\}q$ 也必然为真。

11. 假定程序断言 $p\{S\}q_0$ 和蕴含式 $q_0\rightarrow q_1$ 都为真。证明 $p\{S\}q_1$ 也必然为真。

12. 下面的程序计算商数和余数。

```
r := a
q := 0
while r≥d
begin
    r := r-d
    q := q+1
end
```

验证它是相对于初始断言“a 和 d 都是正整数”和终结断言“q 和 r 是使得 $a=dq+r$ 和 $0\leqslant r<d$ 的整数”为正确的。

13. 用循环不变量去验证欧几里得算法(2.5 节算法 6)是相对于初始断言“a 和 b 都是正整数”和终结断言“$x=\gcd(a, b)$”为部分正确的。

关键术语和结果

术语

序列：以整数集合的子集合作为定义域的函数。

几何序列：形如 a，ar，ar^2，…的序列，其中 a 和 r 都是实数。

等差序列：形如 a，$a+d$，$a+2d$，…的序列，其中 a 和 d 都是实数。

串：有穷序列。

空串：长度为 0 的串。

$\sum_{i=1}^{n}$：和 $a_1+a_2+\cdots+a_n$。

$\prod_{i=1}^{n}$：积 $a_1a_2\cdots a_n$。

可数集：有穷集或可与正整数集合一一对应的集合。

不可数集：不是可数的集合。

康托尔对角化论证：可用来证明实数集合是不可数的证明技术。

数学归纳法原理：命题：若 $P(1)$为真而且 $\forall k[P(k)\rightarrow P(k+1)]$为真，则 $\forall nP(n)$为真。

基础步骤：在 $\forall nP(n)$的数学归纳法证明中对 $P(1)$的证明。

归纳步骤：在 $\forall nP(n)$的数学归纳法证明中对 $P(k)\rightarrow P(k+1)$的证明。

强归纳法：命题：若 $P(1)$为真而且 $\forall k[(P(1)\wedge\cdots\wedge P(k))\rightarrow P(k+1)]$为真，则 $\forall nP(n)$为真。

良序性：非负整数的每个非空集合都有最小元。

函数的递归定义：规定一组初始的函数值以及从较小整数处的函数值获得较大整数处的函数值的规则。

集合的递归定义：规定集合里的一组初始元素以及从已知属于集合的元素获得其他元素的规则。

结构归纳法：证明关于递归定义的集合的结果的技术。

递归算法：通过把问题归约到带有较小输入的同样问题而进行的算法。

归并排序：排序一个表的排序算法，它把一个表分成两个表，对得出的两个表各自进行排序，并且把结果归并成一个有序表。

迭代：基于反复利用循环里的操作的过程。

程序正确性：对过程总是产生正确结果的验证。

循环不变量：在循环的每次执行期间都保持为真的性质。

初始断言：规定程序的输入值所具有的性质的命题。

终结断言：规定若程序正确地工作则输出值所应当具有的性质的命题。

结果

停机问题：不存在这样的过程，当给定一个程序和这个程序的输入时，这个过程确定当给定该输入时该程序是否停机。

几何级数各项之和：当 $r \neq 1$ 时，$\sum_{j=0}^{n} ar^j = (ar^{n+1} - a)/(r-1)$。

有理数集合是可数的。

实数集合是不可数的。

复习题

1. 描述一下形成猜想和解决猜想的过程，并且用下面的例子演示这个过程：寻找形如 a^n-1 的素数，其中 a 和 n 都是整数。
2. 证明：奇整数集合是可数的。
3. 给出一个不可数集合的例子。
4. a)能否用数学归纳法原理来求出一个序列的前 n 项之和的公式？
   b)能否用数学归纳法原理来判定一个序列的前 n 项之和的给定公式是否正确？
   c)求出前 n 个正偶数之和的公式，并且用数学归纳法证明它。
5. a)对哪些正整数 n 来说 $11n+17 \leqslant 2^n$ 为真？
   b)用数学归纳法来证明你在(a)里所做的猜想。
6. a)仅用5分和9分的邮票，可以组成哪些数量的邮资？
   b)用数学归纳法证明你所做的猜想。
   c)用数学归纳法第二原理证明你所做的猜想。
   d)找出与你在(b)和(c)里所给出的证明不同的、对你的猜想的证明。
7. 给出使用强归纳法的三个不同的证明例子。
8. a)叙述正整数集合的良序性。
   b)利用这个性质证明：每个正整数都可以写成素数之积。
9. a)解释一下为什么若通过规定 $f(1)$ 以及从 $f(n-1)$ 求出 $f(n)$ 的规则来递归地定义一个函数，则这个函数是良定义的。
   b)给出函数 $f(n)=(n+1)!$ 的递归定义。
10. a)给出斐波那契数的递归定义。
    b)证明：每当 $n \geqslant 3$ 时，就有 $f_n > \alpha^{n-2}$，其中 f_n 是斐波那契序列的第 n 项而 $\alpha=(1+\sqrt{5})/2$。
11. a)解释一下为什么若通过规定 a_1 和 a_2 以及从 a_1，a_2，…，a_{n-1} ($n=3$，4，5，…)，求 a_n 的规则来递归地定义一个序列，则这个序列是良定义的。
    b)若 $a_1=1$，$a_2=2$，$a_n=a_{n-1}+a_{n-2}+\cdots+a_1$，$n=3$，4，5，…，试求出 a_n 的值。
12. 给出两个例子说明对由元素和运算组成的不同集合来说，如何递归地定义合式公式。
13. a)给出字符串长度的递归定义。
    b)用(a)的递归定义来证明 $l(xy)=l(x)+l(y)$。
14. a)什么是递归算法？
    b)描述计算序列里 n 个数之和的递归算法。
15. 描述计算两个正整数的最大公因子的递归算法。
16. a)描述归并排序算法。
    b)用归并排序算法把表4，10，1，5，3，8，7，2，6，9排成升序。
    c)给出归并排序使用的比较次数的大 O 估计。
17. a)测试一个计算机程序，看看对某些输入值来说它是否产生了正确的输出，是否这样就验证了这个程序总是产生正确的输出？
    b)证明了一个计算机程序是相对于初始断言和终结断言为部分正确的，是否这样就验证了这个程序总是产生正确的输出？若不是，则还需要证明什么别的东西？

18. 可以用什么技术来证明长的计算机程序是相对于初始断言和终结断言为部分正确的?
19. 什么是循环不变量? 如何使用循环不变量?

补充练习

1. 证明：$2^{2^5}+1$ 是合数。
2. a)形成关于哪些形如 a^n+1 的整数必定是合数的猜想，其中 $a\in\mathbf{Z}^+$ 且 $n\in\mathbf{Z}^+$。
   b)证明这个猜想。
3. 证明：$\sqrt[3]{7}$是无理数。
4. 证明：$\sqrt{5}$是无理数。
5. 证明：$x^4+y^4=1000$ 没有整数解。
6. 证明当 n 不能被 5 整除时，n^4-1 能被 5 整除。使用分情形证明，分四种情形——当你把一个不能被 5 整除的整数除以 5 时，每一种非零余数对应一种情形。
7. 证明或反驳：$|x-y|\geqslant|x|-|y|$，x 和 y 是实数。

*8. 把尤兰(Ulam)数定义成：$u_1=1$ 和 $u_2=2$。另外，在确定了小于 n 的整数是否为尤兰数之后，若 n 可以唯一地写成两个不同的尤兰数之和，则让 n 等于下一个尤兰数。注意 $u_3=3$、$u_4=4$、$u_5=6$ 和 $u_6=8$。
   a)求出前 20 个尤兰数。
   b)证明：存在无穷多个尤兰数。

9. 给出构造性的证明:存在多项式 $P(x)$ 使得 $P(x_1)=y_1,P(x_2)=y_2,\cdots,P(x_n)=y_n$,其中 $x_1,\cdots,x_n,y_1,\cdots,y_n$ 都是实数且 x_i 是各不相同的。$\left[\text{提示:设 } P(x)=\sum_{i=1}^{n}\left(\prod_{i\neq j}\frac{x-x_j}{x_i-x_j}\right)y_i\text{。}\right]$
10. 证明:每当 n 是正整数时,就有 $1^3+3^3+5^3+\cdots+(2n+1)^3=(n+1)^2(2n^2+4n+1)$。
11. 证明:每当 n 是正整数时,就有 $1\cdot 2^0+2\cdot 2^1+3\cdot 2^2+\cdots+n\cdot 2^{n-1}=(n-1)\cdot 2^n+1$。
12. 证明:每当 n 是正整数时,就有
$$\frac{1}{1\cdot 3}+\frac{1}{3\cdot 5}+\cdots+\frac{1}{(2n-1)(2n+1)}=\frac{n}{2n+1}$$
13. 证明:每当 n 是正整数时,就有
$$\frac{1}{1\cdot 4}+\frac{1}{4\cdot 7}+\cdots+\frac{1}{(3n-2)(3n+1)}=\frac{n}{3n+1}$$
14. 证明:每当 n 是大于 4 的正整数时,就有 $2^n>n^2+n$。
15. 用数学归纳法证明:每当 n 是大于 9 的正整数时,就有 $2^n>n^3$。
16. 求出整数 N,使得每当 n 大于 N 时,就有 $2^n>n^4$。用数学归纳法证明你的结果是正确的。
17. 用数学归纳法证明:每当 n 是正整数时,就有 $a-b$ 是 a^n-b^n 的因子。
18. 用数学归纳法证明:每当 n 是非负整数时,就有 9 整除 $n^3+(n+1)^3+(n+2)^3$。
19. 用数学归纳法证明算术序列各项之和的这个公式:
$$a+(a+d)+\cdots+(a+nd)=(n+1)(2a+nd)/2$$
20. 假定对 $j=1,2,\cdots,n$ 来说 $a_j\equiv b_j(\bmod m)$。用数学归纳法证明:
    a) $\sum_{j=1}^{n}a_j\equiv\sum_{j=1}^{n}b_j(\bmod m)$
    b) $\prod_{j=1}^{n}a_j\equiv\prod_{j=1}^{n}b_j(\bmod m)$
21. 证明:若 n 是正整数,则
$$\sum_{k=1}^{n}\frac{k+4}{k(k+1)(k+2)}=\frac{n(3n+7)}{2(n+1)(n+2)}$$
22. 对于哪些正整数 n 来说 $n+6<(n^2-8n)/16$?用数学归纳法证明你的答案。

23.（本题需要微积分知识）假设 $f(x)=e^x$ 并且 $g(x)=xe^x$。用数学归纳法以及乘积求导规则和 $f'(x)=e^x$ 的事实来证明：每当 n 是正整数时 $g^{(n)}(x)=(x+n)e^x$。

24.（本题需要微积分知识）假设 $f(x)=e^x$ 并且 $g(x)=e^{cx}$，其中 c 是常数。用数学归纳法以及复合求导规则和 $f'(x)=e^x$ 的事实来证明：每当 n 是正整数时 $g^{(n)}=c^n e^{cx}$。

*25. 确定哪些斐波那契数是偶数，用数学归纳法的一种形式来证明你的猜想。

*26. 确定哪些斐波那契数能被 3 整除，用数学归纳法的一种形式来证明你的猜想。

*27. 证明：对所有非负整数 n 来说，$f_k f_n+f_{k+1}f_{n+1}=f_{n+k+1}$，其中 k 是非负整数而且 f_i 表示第 i 个斐波那契数。

卢卡斯(Lucas) *数*的序列定义成：$l_0=2, l_1=1$ 以及对 $n=2,3,4,\cdots$ 来说 $l_n=l_{n-1}+l_{n-2}$。

28. 证明：每当 n 是正整数时，就有 $f_n+f_{n+2}=l_{n+1}$，其中 f_i 和 l_i 分别是第 i 个斐波那契数和第 i 个卢卡斯数。

29. 证明：每当 n 是非负整数而且 l_i 是第 i 个卢卡斯数时，就有 $l_0^2+l_1^2+\cdots+l_n^2=l_n l_{n+1}+2$。

*30. 用数学归纳法证明：任意 n 个连续正整数之积能被 $n!$ 整除。[提示：利用恒等式 $m(m+1)\cdots(m+n-1)/n!=(m-1)m(m+1)\cdots(m+n-2)/n!+m(m+1)\cdots(m+n-2)/(n-1)!$。]

31. 用数学归纳法证明：每当 n 是正整数时，就有 $(\cos x+i\sin x)^n=\cos nx+i\sin nx$。[提示：利用恒等式 $\cos(a+b)=\cos a\cos b-\sin a\sin b$ 和 $\sin(a+b)=\sin a\cos b+\cos a\sin b$。]

*32. 用数学归纳法证明：每当 n 是正整数而且 $\sin(x/2)\neq 0$ 时，就有

$$\sum_{j=1}^{n}\cos jx=\cos[(n+1)x/2]\sin(nx/2)\sin(x/2)$$

33. 用数学归纳法证明：对于每个正整数 n 来说

$$\sum_{j=1}^{n}j^2 2^j=n^2 2^{n+1}-n2^{n+2}+3\cdot 2^{n+1}-6$$

34.（本题要求微积分知识）假设序列 $x_1,x_2,\cdots,x_n,\cdots$ 递归地定义成 $x_1=0$ 和 $x_{n+1}=\sqrt{x_n+6}$。

a)用数学归纳法证明：$x_1<x_2<\cdots<x_n<\cdots$，即序列$\{x_n\}$是单调增的。

b)用数学归纳法证明：对于 $n=1,2,\cdots$来说 $x_n<3$。

c)证明：$\lim_{n\to\infty}x_n=3$。

35. 用数学归纳法证明：若 n 个人站成一队，其中 n 是正整数，并且若该队中第一个人是女人，最后一个人是男人，则队中某处有一个女人直接站在一个男人前面。

*36. 假设在一个国家中有直达的单行道路连接每一对城市。用数学归纳法证明：存在一个城市，从其他每个城市都可以直达这个城市，或者恰好经由一个其他城市而到达这个城市。

37. 用数学归纳法证明：当 n 个圆周把平面分成区域时，这些区域可以用两种颜色着色，使得具有共同边界的区域都不染成同样的颜色。

*38. 假设有足够的燃料让环行赛道上一组汽车中的一辆跑完一圈。用数学归纳法证明：在这组汽车中存在一辆汽车，当它沿着赛道前进时，可以通过从其他汽车获得加油来跑完一圈。

39. 证明：如果 n 是正整数，则 $\sum_{j=1}^{n}(2j-1)\left(\sum_{k=j}^{n}1/k\right)=n(n+1)/2$。

40. *单位分数*或*埃及分数*是形如 $1/n$ 的分数，其中 n 是正整数。在本题中，将用强归纳法证明：可以用贪心算法把每个满足 $0<p/q<1$ 的有理数 p/q 表达成不同的单位分数之和。在算法的每一步，求出最小的正整数 n 使得这个和可以加入 $1/n$ 而不超过 p/q。例如，为了表达 5/7，从 1/2 这个和开始。由于$5/7-1/2=3/14$，所以把 1/5 加入这个和，因为 5 是最小的正整数 k 使得 $1/k<3/14$。由于 $3/14-1/5=1/70$，所以算法终止，证明 $5/7=1/2+1/5+1/70$。设 $T(p)$是命题：对于所有满足 $0<p/q<1$ 的有理数 p/q 来说这个算法终止。通过证明对于所有正整数 p 来说 $T(p)$为真，将证明这个算法总是终止。

a)证明基础步骤 $T(1)$成立。

b)假设对于满足 $k<p$ 的正整数 k 来说 $T(k)$成立。换句话说，假设对于所有有理数 k/r 来说算法终止，其中 $1\leqslant k<p$。证明：如果从 p/q 开始并且算法第一步选择分数 $1/n$，则 $p/q=p'/q'+1/n$，其中 $p'=np-q$ 且 $q'=nq$。在考虑 $p/q=1/n$ 的情形之后，用归纳假设证明：当贪心算法从 p'/q' 开始时，这个算法总是终止，从而完成归纳步骤。

麦卡锡[⊖]91 函数（人工智能的奠基人之一 John McCarthy 所定义）定义成：对所有正整数 n 来说应用规则

$$M(n)=\begin{cases} n-10 & 若\ n>100 \\ M(M(n+11)) & 若\ n\leqslant 100 \end{cases}$$

41. 通过连续地使用 $M(n)$ 的定义规则来求出

a) $M(102)$　　b) $M(101)$　　c) $M(99)$　　d) $M(97)$　　e) $M(87)$　　f) $M(76)$

**42. 证明：函数 $M(n)$ 是良定义的从正整数集合到正整数集合的函数。[提示：证明对所有满足 $n\leqslant 101$ 的正整数 n 来说都有 $M(n)=91$。]

43. 下述的证明：每当 n 是正整数时就有

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\cdots+\frac{1}{(n-1)n}=\frac{3}{2}-\frac{1}{n}$$

是否正确？为你的答案给出理由。

基础步骤：当 $n=1$ 时结果为真，因为

$$\frac{1}{1\cdot 2}=\frac{3}{2}-\frac{1}{1}$$

归纳步骤：假定对 n 来说结果为真。则

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\cdots+\frac{1}{(n-1)n}+\frac{1}{n(n+1)}=\frac{3}{2}-\frac{1}{n}+\left(\frac{1}{n}-\frac{1}{n+1}\right)=\frac{3}{2}-\frac{1}{n+1}$$

因此，若对 n 来说结果为真，则对 $n+1$ 来说结果为真。证毕。

*44. 一种拼图玩具是通过把可以拼在一起的图片合并成块来完成的。每次当把一个图片加入一个块时，或者当把两个块合并时，就算作一步。用数学归纳法的第二种形式证明：无论如何完成各步，组装含有 n 个图片的拼图玩具都恰好需要 $n-1$ 步。

*45. 证明：若在 n 个圆中每两个都恰好相交于两点，而任意三个都没有公共点，则这些圆把平面划分成 n^2-n+2 个区域。

*46. 证明：若在 n 个平面中任意三个都有公共点，而任意四个都没有公共点，则这些平面把三维空间划分成 $(n^3+5n+6)/6$ 个区域。

*47. 用良序性证明：$\sqrt{2}$ 是无理数。（提示：假定 $\sqrt{2}$ 是有理数。证明形如 $b\sqrt{2}$ 的正整数组成的集合有最小元 a。然后证明 $a\sqrt{2}-a$ 是具有这种形式的更小的正整数。）

48. 若一个集合的每个非空子集合都有最小元，则这个集合是*良序的*。判断下面的每个集合是否良序的。

a) 整数集合　　b) 大于 -100 的整数的集合

c) 正有理数集合　　d) 分母小于 100 的正有理数的集合

*49. 证明：若把数学归纳法原理作为一条公理，则可以证明良序性。

*50. 证明：数学归纳法的第一和第二原理是等价的；即从一个原理是有效的就可以证明另外一个原理是有效的。

51. a) 证明：若 $a_1, a_2, \cdots, a_n$ 都是正整数，则 $\gcd(a_1, a_2, \cdots, a_{n-1}, a_n)=\gcd(a_1, a_2, \cdots, a_{n-2}, \gcd(a_{n-1}, a_n))$。



⊖ **约翰·麦卡锡（John McCarthy，1927 年生）**麦卡锡出生在波士顿。在波士顿和洛杉矶长大。他在本科和研究生阶段学习的都是数学。1948 年他从加州理工学院获得学士学位并且在 1951 年从普林斯顿获得博士学位。从普林斯顿毕业后，麦卡锡在普林斯顿大学、斯坦福大学、达特茅斯和麻省理工学院任职。从 1962 年到 1994 年他一直在斯坦福任职，是那里的荣誉教授。在斯坦福，他是人工智能实验室的主任，担任工程学院的名誉院长，并且是胡佛学院的资深院士。

麦卡锡是人工智能研究的先锋，“人工智能”就是他在 1955 年发明的术语。他致力于关于智能的计算机行为所需要的推理和信息需求的问题。麦卡锡是设计分时计算机系统的首批计算机科学家之一。他开发了 LISP，这是用符号表达式来计算的一种程序设计语言。他在用逻辑来验证计算机程序的正确性方面起到了重要作用。麦卡锡还致力于研究计算机技术的社会影响问题。他目前正在致力于在情况不是错综复杂的假设下人和计算机如何形成猜想的问题。麦卡锡是人类可持续性发展的倡导者并且是关于人类未来的乐观者。他还编写科幻故事。他最近的一些著作探索了我们的世界是由更高的力量所编写的计算机程序的可能性。

麦卡锡所获得过的国际奖励有：计算机学会的图灵奖，国际人工智能会议的杰出研究工作奖，京都奖和美国国家科学奖章。

b)利用(a)和欧几里得算法得出一个计算 n 个正整数的最大公因子的递归算法。

*52. 描述一个递归算法，把 n 个正整数的最大公因子表示成这些整数的线性组合。

53. 求出 $f(n)$ 的显式公式，其中 $f(1)=1$ 而且若 $n\geqslant 2$ 则 $f(n)=f(n-1)+2n-1$。用数学归纳法证明你的结果。

**54. 给出由所含有的 0 是 1 的两倍的位串所组成的集合的递归定义。

55. 设 S 是位串的集合，它递归地定义成：$\lambda\in S$，以及若 $x\in S$，则 $0x\in S$，$x1\in S$，其中 λ 是空串。

a)求出 S 中所有长度不超过 5 的串。

b)给出对 S 中元素的显式描述。

56. 设 S 是字符串的集合，它递归地定义成：$abc\in S$，$bac\in S$，$acb\in S$，以及若 $x\in S$ 则 $abcx\in S$，$abxc\in S$，$axbc\in S$ 和 $xabc\in S$。

a)求出 S 中长度为 8 或更短的所有串。 b)证明：S 中的每个元素都有能被 3 整除的长度。

由所有平衡的括号串组成的集合递归地定义成：$\lambda\in B$，其中 λ 是空串；若 x，$y\in B$，则 $(x)\in B$，$xy\in B$。

57. 证明：(()())是平衡的括号串而(()))不是平衡的括号串。

58. 求出所有恰好带 6 个符号的平衡的括号串。

59. 求出所有带 4 个或更少符号的平衡的括号串。

60. 用归纳法证明：若 x 是平衡的括号串，则在 x 里左括号的个数等于右括号的个数。

在括号串的集合上定义函数 N 为：

$$N(\lambda)=0,\ N(()=1,\ N())=-1,$$
$$N(uv)=N(u)+N(v),$$

其中 λ 是空串，u 和 v 都是串。可以证明 N 是良定义的。

61. 求出

a) $N(())$ b) $N()))())(()$ c) $N((()(())$ d) $N(()((()))(()))$

**62. 证明：括号串 w 是平衡的当且仅当 $N(w)=0$，而且每当 u 是 w 的前缀(即 $w=uv$)时，就有 $N(u)\geqslant 0$。

*63. 给出一个求所有包含 n 个或更少符号的平衡的括号串的递归算法。

64. 根据下述的事实：若 $a=b$ 则 $\gcd(a, b)=a$，若 a 和 b 都是偶数则 $\gcd(a, b)=2\gcd(a/2, b/2)$，若 a 是偶数而 b 是奇数则 $\gcd(a, b)=\gcd(b-a, b)$，给出一个求满足 $a\leqslant b$ 的两个非负整数 a 和 b 的最大公因子的递归算法。

65. 验证程序段

```
if  x>y  then
    x := y
```

是相对于初始断言 T 和终结断言 $x\leqslant y$ 为正确的。

*66. 发展一条验证递归程序的推理规则，并用它去验证在 3.5 节给出的计算阶乘的递归程序。

67. 设计求整数表中整数 0 出现次数的递归算法。

练习 68～75 处理某些不寻常的序列，这些序列非正式地称为*自生成序列*，它们是用简单的递归关系或规则产生的。具体地说，练习 68～75 处理序列 $\{a(n)\}$，它定义成：对 $n\geqslant 1$ 来说 $a(n)=n-a(a(n-1))$，而且 $a(0)=0$。(这个序列以及在练习 72 和练习 73 里的序列，都是在道格拉斯·霍夫斯塔德的奇妙的书《歌德尔、埃舍尔、巴赫》[Ho99]中定义的)。

68. 求出在本题前面的说明中定义的序列 $\{a(n)\}$ 的前 10 项。

*69. 证明：这个序列是良定义的。即证明对所有非负整数 n 来说 $a(n)$ 是唯一地定义的。

**70. 证明：$a(n)=\lfloor (n+1)\mu\rfloor$，其中 $\mu=(-1+\sqrt{5})/2$。[提示：首先证明对所有 $n>0$ 来说 $(\mu n-\lfloor \mu n\rfloor)+(\mu^2 n-\lfloor \mu^2 n\rfloor)=1$。然后证明对满足 $0\leqslant\alpha<1$ 和 $\alpha\neq 1-\mu$ 的所有实数 α 来说 $\lfloor (1+\mu)(1-\alpha)\rfloor+\lfloor \alpha+\mu\rfloor=1$，分别考虑 $0\leqslant\alpha<1-\mu$ 和 $1-\mu<\alpha<1$ 的情形。

*71. 利用练习 70 的公式证明：若 $\mu n-\lfloor \mu n\rfloor<1-\mu$ 则 $a(n)=a(n-1)$，否则 $a(n)=a(n-1)+1$。

72. 求出下面每个自生成序列的前 10 项：

a)对 $n\geqslant1$ 来说 $a(n)=n-a(a(a(n-1)))$，而且 $a(0)=0$。

b)对 $n\geqslant1$ 来说 $a(n)=n-a(a(a(a(n-1))))$，而且 $a(0)=0$。

c)对 $n\geqslant3$ 来说 $a(n)=a(n-a(n-1))+a(n-a(n-2))$，而且 $a(1)=1$ 和 $a(2)=1$。

73. 求出序列 $m(n)$和 $f(n)$的前 10 项，它们是用下面的嵌套的递归关系来定义的：对 $n\geqslant1$ 来说 $m(n)=n-f(m(n-1))$，$f(n)=n-m(f(n-1))$，而且 $f(0)=1$ 和 $m(0)=0$。

哥伦布的自生成序列是具有下述性质的、唯一的非减的正整数序列 a_1，a_2，a_3，…：对每个正整数 k 来说这个序列恰好包含 k 的 a_k 次出现。

74. 求出哥伦布的自生成序列的前 20 项。

*75. 证明：若 $f(n)$ 是使得 $a_m=n$ 的最大整数 m，其中 a_m 是哥伦布的自生成序列的第 m 项，则 $f(n)=\sum_{k=1}^{n}a_k$ 而且 $f(f(n))=\sum_{k=1}^{n}ka_k$。

G·H·哈代[1]引入的对数指数函数集合 L 是使得下面性质成立的最小的函数集合：

- 函数 $f(n)=a$ 属于 L，其中 a 是一个正实数；
- 函数 $f(n)=n$ 属于 L；
- 若函数 $f(n)$和 $g(n)$都属于 L，则 $f(n)-g(n)$属于 L；
- 若函数 $f(n)$属于 L，则 $e^{f(n)}$ 属于 L；
- 若 $f(n)$属于 L 而且存在整数 N 使得对 $n\geqslant N$ 来说 $f(n)>0$(这意味着 f 是逐渐正的)，则 $\ln f(n)$属于 L，其中 $\ln x$ 表示 x 的对数。

哈代证明，对每个不恒等于 0 的对数指数函数来说，$f(n)$或者是逐渐正的，或者是逐渐负的。他证明若 $f(n)$和 $g(n)$都属于 L，则或者 $f(n)$是 $o(g(n))$，或者 $g(n)$是 $o(f(n))$，或者 $f(n)$和 $g(n)$有同样的阶。

76. 证明：若 $f(n)$和 $g(n)$都属于 L，则 $f(n)+g(n)$属于 L。

*77. 证明：若 $f(n)$和 $g(n)$都属于 L 而且都是逐渐正的，则 $f(n)g(n)$和 $f(n)/g(n)$都属于 L，利用每个对数



[1] 葛弗雷·哈罗得·哈代(Godfrey Harold Hardy，1877—1947) 哈代出生在英格兰的色雷的克兰雷，是伊萨克·哈代与索菲亚·哈尔·哈代的两个孩子中的长子。他的父亲是克兰雷中学的地理学与绘图教师，还教过唱歌，踢过足球。他的母亲教授钢琴，而且帮助经营宿舍，为年轻学生提供膳宿。哈代的父母亲对子女的教育全力以赴。哈代 2 岁时写出上百万的数字，显示了他的数字能力。他受教于一位私人数学教师而没有上克兰雷中学的正规课程。他 13 岁时转到温彻斯特学校并被这所私立高中授予奖学金。他学习成绩优秀，对数学显示出强烈的兴趣。他在 1896 年获得奖学金进入剑桥的三一学院，在学期间几次获奖，于 1899 年毕业。

哈代从 1906 年到 1919 年在剑桥大学三一学院担任数学讲师职务，然后被任命为牛津的几何学的苏里文教授。他在剑桥担任纯数学的萨德雷里安教授，一直到 1942 年退休为止。他是一个纯数学家并对数学持有卓尔不群的观点，希望他的研究永远不被应用。具有讽刺意味的是，他最出名的成果也许是作为发明人之一，发明了预言遗传模式的哈代—温伯格法则。在给《科学》杂志的一封信中，他用简单的代数想法指出了一篇遗传学论文的错误。哈代的研究范围主要在数论和函数论方面，研究黎曼 zeta 函数、傅里叶级数、素数分布这样的课题。他对许多重要问题作出了重要的贡献，比如关于把正整数表示成 k 次幂之和的华林问题，以及把奇数表示成 3 个素数之和的问题。哈代还因为与剑桥的同事约翰·E·利特伍德和著名的印度数学奇才斯里尼瓦萨·拉马努金的合作而为人们所记忆。他与利特伍德合写的论文有 100 多篇，他们的合作引出这样的笑话：在当时只有 3 个重要的英国数学家：哈代、利特伍德和哈代—利特伍德，甚至有人认为哈代发明了虚构的人物利特伍德，因为在剑桥以外很少看见利特伍德。哈代能从拉马努金寄给他的非常规的但是极有创造性的结果中，识别出拉马努金的天才，而这点其他数学家却没能做到。哈代把拉马努金带到剑桥，合作写出重要的联名文章，证明了关于整数的分拆数的新结果。哈代对数学教育很感兴趣，他的著作《纯数学教程》在 20 世纪的前 50 年里深深地影响了数学的本科教学。哈代还写过《一个数学家的辩解》，在这本书里他回答了这样的问题：一个人是否值得把一生都献给数学研究。哈代在这本书给出了关于数学是什么和数学做什么的观点。

哈代对体育有着强烈的兴趣。他是狂热的板球迷，对分数盯得很紧。他有一个特别的习惯，不喜欢拍照(只有 5 张照片是为人所知的)，也不喜欢镜子，他一进入旅馆房间就立即用毛巾把镜子盖上。

指数函数都是或者逐渐正，或者逐渐负，或者恒等于 0 的事实来证明，前面的结果蕴含着任何两个不恒等于 0 的属于 L 的函数之积与商都属于 L。

78. 利用练习 76 和练习 77 证明：每个带实系数的多项式 $f(n)=a_m n^m+a_{m-1}n^{m-1}+\cdots+a_0$ 属于 L。
79. 证明：若 $f(n)$属于 L，则 $\sqrt{f(n)}$属于 L。
80. 证明：函数 $e^{\sqrt{n}\sqrt{\ln n \ln \ln n}}$属于 L。

计算机题目

按下述给定的输入和输出写程序。

1. 给定序列项 $a_1,a_2,\cdots,a_n$,求 $\sum_{j=1}^{n} a_j$ 和 $\prod_{j=1}^{n} a_j$。
2. 给定几何级数 $a,ar,ar^2,\cdots,ar^n$,求出它的各项之和。
3. 给定非负整数 n,求出 n 个最小的正整数之和。

**4. 给定去掉一个格子的 $2^n\times 2^n$ 棋盘,用 L 形状的拼片构造出这个棋盘。

**5. 对含有变量 x、y 和 z 以及运算符{+, *, /, -}的表达式来说,生成所有的带有 n 个或更少符号的合式公式。

6. 生成所有带有 n 个或更少符号的命题的合式公式,其中每个符号是 **T、**F**,命题变量 p 和 q 之一或{¬, ∨, ∧, →, ↔}里的一个运算符。

7. 给定一个字符串,求出它的倒置。
8. 给定实数 a 和非负整数 n,用递归求 a^n。
9. 给定实数 a 和非负整数 n,用递归求 a^{2^n}。



斯里尼瓦萨·拉马努金(Srinivasa Ramanujan, 1887—1920) 拉马努金是数学奇才，诞生和成长在靠近马德拉斯城的南印度。他的父亲是布店的店员，母亲在当地的寺庙里唱歌来增加家庭收入。拉马努金在当地的英语学校学习时显示出在数学方面的天赋和兴趣。在 13 岁时他就掌握了大学数学课程。15 岁时，一个大学生借给他一本《纯数学提要》。拉马努金决心验证这本书里只有叙述而没有证明和解释的 6000 多个结论，写在纸片上，后来收集成笔记。他 1904 年从高中毕业，赢得马德拉斯大学的奖学金。但由于注册的是美术专业，而他只顾学习数学，结果丢了奖学金。1904 年到 1907 年他在大学里 4 次考试不及格，只在数学上成绩优异。在这段时间里他在笔记本里写下了创造性的结果，有时重新发现了已经发表过的结果，有时则有了新的发现。

由于没有大学学位，拉马努金难以找到好的工作。为了谋生，他不得不依赖朋友们的施舍。他教过学生数学，但是他非同寻常的思考方式和不遵循教学大纲却引来了麻烦。他在 1909 年接受了包办婚姻，与比他小 9 岁的年轻女子结婚。为了养活自己和妻子，他搬到马德拉斯并且找了个工作。他向可能雇佣他的人出示他的记有数学结果的笔记本，但是却惹恼了他们。不过，总统学院的一个教授看出他的天才并且在一段时期内支持过他，在 1912 年他找到了一份薪水微薄的会计工作。

拉马努金在这段时间里继续他的数学工作，并且在 1910 年在一份印度期刊上发表了他的第一篇论文。他认识到他的工作超出了印度数学家的理解范围，故决定给最重要的英国数学家写信。第一个收到信的数学家拒绝了他寻求帮助的要求。1913 年 1 月他写信给 G·H·哈代，哈代也曾想拒绝，但是信中的数学命题却让哈代感到迷惑，尽管没有证明。他决定在同事与合作者 J·E·利特伍德的帮助下仔细地检查它们。经过仔细研究，他们断定拉马努金可能是个天才，因为他的那些命题“只能是一流数学家写出来的；它们一定是对的。假如它们不对，那么就没有人具有发明它们的想像力。”

哈代为拉马努金安排了奖学金，在 1914 年把他带到英格兰。哈代亲自教他数学分析。他们合作了 5 年，证明了关于整数的分拆数的重要定理。在这段时间里，拉马努金对数论作出了重要贡献，还在连分数、无穷级数以及椭圆曲线方面做了大量工作。拉马努金对特定类型的函数和级数具有惊人的洞察力，但是他关于素数的所谓定理却常常是错的，暴露出他对什么是正确证明的模糊观点。他是被任命为皇家协会成员的最年轻的成员之一。不幸的是，在 1917 年拉马努金得了重病。当时，他被认为不适应英国的气候并且感染了肺结核。现在看来是维生素缺乏症，起因是拉马努金严格的素食主义和战时英国的食物匮乏。他在 1919 年回到印度，即便卧床时，也继续研究数学。他是教徒并且认为他的数学才能来自家庭的守护神——拉马吉里。他认为数学与宗教是有联系的。他说：“一个等式对我来说毫无意义，除非它表达了神的思想。”1920 年 4 月他短暂的生命走到了终点，当时他 32 岁。拉马努金留下了写有未发表结果的几个笔记本。这些笔记本里的结果说明了拉马努金的洞察力，但都是简洁的勾勒。一些数学家花了许多年的时间来解释和证实这些笔记本里的结果。

*10. 给定实数 a 和非负整数 n，利用 n 的二进制展开式和计算 a^{2^k} 的递归算法来求 a^n。

11. 给定两个不全为零的整数，用递归求它们的最大公因子。

12. 给定整数的列表和元素 x，用线性搜索的递归实现求 x 在这个列表中的位置。

13. 给定整数的列表和元素 x，用二叉搜索的递归实现求 x 在这个列表中的位置。

14. 给定非负整数 n，用迭代来求第 n 个斐波那契数。

15. 给定非负整数 n，用递归来求第 n 个斐波那契数。

16. 给定一个正整数，求出这个整数的划分的数目。（参见 3.4 节练习 47。）

17. 给定正整数 m 和 n，求出阿克曼函数在 (m,n) 处的值 $A(m,n)$。（参见 3.4 节练习 48 前面的说明。）

18. 给定 n 个整数的列表，用归并排序给这些整数排序。

计算和研究

用计算程序或你已经编写的程序做下面的练习。

1. 对 $n \leqslant 10\,000$ 验证哥德巴赫猜想，它断言每个正偶数 n 都是两个素数之和。
2. 对满足 $n \leqslant 20$ 的所有正整数 n，求出 $n!+1$ 的最小素因子。
3. 对满足 $n \leqslant 10$ 的每个正整数 n，求出最小的 n 个连续的合数。
4. 找出尽可能多的孪生素数。
5. 对于尽可能多的正整数验证 $3x+1$ 猜想。
6. 让 $n!$ 具有不超过 100 位十进制数字和不超过 1000 位十进制数字的 n 的最大值是多少？
7. 确定哪些斐波那契数能被 5 整除、哪些能被 7 整除、哪些能被 11 整除。证明你的猜想是正确的。
8. 用 L 形状的碎片构造出去掉一个格子的 16×16、32×32 和 64×64 的棋盘。
9. 探索用 L 形状的碎片可以完全地覆盖哪些 $m\times n$ 棋盘。能否形成关于这个问题的猜想？
10. 阿克曼函数的哪些值是足够小的使得能够计算出它们？
11. 比较一下递归地计算斐波那契数与迭代地计算它们所需要的运算次数或时间。

写作题目

利用本书以外的资料，针对下面的内容写出短文。

1. 简要解释一下用来证明目前已知的关于哥德巴赫猜想的结果的各种方法。
2. 描述一下已经形成的关于不超过实数 x 的孪生素数的个数的猜想。这些猜想是在什么基础上形成的？
3. 描述一下 3.1 节所述之外的某个数论中的未解决问题。
4. 解释一下发现《整数序列百科全书》有用的不同方式。另外，描述一下这本百科全书中几个较为不寻常的序列以及这些序列的起源。
5. 定义一下最近发明的 EKG 序列，并且描述一下它的一些性质和关于它的未解决问题。
6. 查阅一下超越数的定义，解释一下如何证明这种数存在以及如何构造这种数。可以证明的著名的超越数有哪些？哪些著名的数还不知道是否为超越数？
7. 描述一下数学归纳法的起源。谁是第一批使用它的人？他们在哪个问题上用到了它？
8. 在过去的 25 年里，基于大量的计算机计算，已经证明了几个重要的定理。讨论一下这样的证明的有效性，并且描述一下围绕着基于计算机计算的证明的争论。
9. 3.3 节的练习里所使用的 L 形状的碎片是 Golomb 在 1954 年引入的多元多米诺骨牌的一些例子。描述一下关于用多元多米诺骨牌铺满棋盘的一些问题和有关的结果。
10. 描述一下在递归定义的理论里以及在集合合并算法的复杂性分析里对阿克曼函数的使用情况。
11. 讨论一些用来证明程序正确性的方法，并且将它们与 3.6 节所描述的霍尔方法作一下比较。
12. 解释一下如何扩充程序正确性的思想和概念去证明操作系统是安全的。

第4章 计　　数

组合数学这一研究个体安排的学科，是离散数学的重要部分。早在17世纪就开始了这类课题的研究，当时在赌博游戏的研究中出现了组合问题。枚举，具有确定性质的个体的计数，是组合数学的一个重要的部分。我们必须对个体计数来求解许多不同类型的问题。例如，用计数确定算法的复杂性。计数也用于确定是否存在着能够充分满足需求的电话号码或因特网地址。此外，计数技术也广泛用于计算事件的概率。

4.1节将要研究的基本计数规则可以求解各种各样的问题。例如，可以用这些规则来计数美国各种不同可能的电话号码，计算机系统中允许使用的密码，以及在比赛结束时赛跑运动员的不同的名次。另一个重要的组合工具是鸽巢原理，将放在4.2节研究。这个原理指出，当把物体放在盒子里时，若物体比盒子多，那么存在一个盒子包含着至少两个物体。例如，我们可以用这个原理证明在15个或者更多的学生中至少有3人出生在相同的星期几。

我们可以用集合中个体的有序或无序安排来描述许多计数问题。这些安排称作排列和组合，在许多的计数问题中都会用到它们。例如，在2000个学生参加的考试竞赛中最终将有100个获胜者被邀请赴宴。我们可以枚举将被邀请的100个学生的可能的组合，以及最终10名获奖者的产生方式。

组合数学的另一个问题涉及到生成某个特定类型的所有排列。这在计算机模拟中通常是很重要的。我们将设计算法来生成各种类型的排列。

4.1 计数的基础

4.1.1 引言

一个计算机系统的密码由6、7或8个字符组成。每个字符必须是数字或字母表中的字母。每个密码必须至少包含一位数字。问有多少个这样的密码？本节将介绍回答这个问题及各种各样的其他计数问题所需要的技术。

数学和计算机科学中存在着计数问题。例如，我们必须为成功的实验结果和所有可能的实验结果计数，以确定离散事件的概率。我们需要对某个算法用到的操作数计数，以便研究它的时间复杂性。

本节将介绍基本的计数方法。这些方法是几乎所有计数技术的基础。

4.1.2 基本的计数原则

Assessment 我们将提出两个基本的计数原则，*乘积法则*和*求和法则*。然后将说明怎样用它们来求解许多不同的计数问题。

当一个过程由独立的任务组成时使用乘积法则。

乘积法则　假定一个过程可以被分解成两个任务。如果完成第一个任务有n_1种方式，在第一个任务完成之后有n_2种方式完成第二个任务，那么完成这个过程有n_1n_2种方式。

Extra Examples 例1至例9显示怎样使用乘积法则。

例1　用一个字母和一个不超过100的正整数给礼堂的座位编号。那么不同编号的座位最多有多少？

解　给一个座位编号的过程由两个任务组成，即从 26 个字母中先选择一个字母分配给这个座位，然后再从 100 个正整数中选择一个整数分配给它。乘积法则表明一个座位可以有 $26 \cdot 100 = 2600$ 种不同的编号方式。因此，不同编号的座位数至多是 2600。■

例 2　某个计算机中心有 32 台微机，每台微机有 24 个端口。问在这个中心里有多少个不同的单机端口？

解　选择一个端口的过程由两个任务组成。首先挑一台微机，然后在这台微机上挑一个端口。因为有 32 种方式选择微机，而不管选择了哪台微机，又有 24 种方式选择端口，由乘积法则存在 $32 \cdot 24 = 768$ 个端口。■

经常会用到推广的乘积法则。假定一个过程由执行任务 T_1，T_2，…，T_m 来完成。如果在完成任务 T_1，T_2，…，T_{i-1}之后用 n_i 种方式来完成 T_i，那么完成这个过程有 $n_1 \cdot n_2 \cdot \cdots \cdot n_m$ 种方式。可以由两个任务的乘积法则通过数学归纳法证明推广的乘积法则(见本节后的练习 56)。

例 3　有多少个不同的 7 位二进制串？

解　每位有两种选择方式，可以是 0 或 1。因此，乘积法则表明总共有 $2^7 = 128$ 个不同的 7 位二进制串。■

— — — — — — —
每位有2个选择

例 4　如果每个车牌由 3 个字母后跟 3 个数字的序列构成(任何字母的序列都允许)，那么有多少个不同的有效的车牌？

解　对 3 个字母中的每个字母有 26 种选择，对 3 个数字中的每个数字有 10 种选择。因此，由乘积法则总共有 $26 \cdot 26 \cdot 26 \cdot 10 \cdot 10 \cdot 10 = 17\,576\,000$ 个可能的车牌。■

— — —　　— — —
每个字母有26种选择　　每个数字有10种选择

例 5　**计数函数**　从一个 n 元集到一个 m 元集存在多少个函数？

解　一个函数对于定义域中 m 个元素中的每个元素都要选择陪域中 n 个元素中的一个元素来对应。因此，由乘积法则存在 $n \cdot n \cdot \cdots \cdot n = n^m$ 个从 m 元集到 n 元集的函数。例如，从一个 3 元集到一个 5 元集存在 5^3 个不同的函数。■

例 6　**计数一对一函数**　从一个 m 元集到一个 n 元集存在多少个一对一函数？

解　首先注意到当 $m > n$ 时没有从 m 元集到 n 元集的一对一函数。现在令 $m \leqslant n$. 假设定义域中的元素是 a_1，a_2，…，a_m。有 n 种方式选择函数在 a_1 的值。因为函数是一对一的，可以有 $n-1$种方式选择函数在 a_2 的值(因为 a_1 用过的值不能再用)。一般地，有 $n-k+1$ 种方式选择函数在 a_k 的值。由乘积法则，从一个 m 元集到一个 n 元集存在着 $n(n-1)(n-2)\cdots(n-m+1)$个一对一函数。例如，从一个 3 元集到一个 5 元集存在 $5 \cdot 4 \cdot 3 = 60$ 个一对一函数。■



例 7　**电话编号计划**　在北美，电话号码的格式是由一个编号计划规定的。一个电话号码由 10 个数字组成，这些数字由一个 3 位的地区代码、一个 3 位的局代码以及一个 4 位的话机代码组成。出于信号的考虑，在一些数字上有某种限制。为了规定允许的格式，令 X 表示可以在 0 到 9 之间任意选取的数字，N 表示可以在 2 到 9 之间选取的数字，而 Y 表示必须取 0 或 1 的数字。下面讨论两个编号计划，分别称为老计划和新计划(老计划是 20 世纪 60 年代使用的，已经被新计划代替了，但目前对新号码需求的迅速增长甚至使得这个新计划也将显得落伍了)。正如将要证明的，新计划允许使用更多的号码。

在老计划中，地区代码、局代码和话机代码的格式分别为 NYX、NNX 和 XXXX，因而电话号码形式为 NYX-NNX-XXXX。在新计划中，这些代码的格式分别为 NXX、NXX 和 XXXX，因而电话号码形式为 NXX-NXX-XXXX。在老计划和新计划下分别可能有多少个不同

的北美电话号码？

解 由乘积法则，格式为 NYX 的地区代码有 $8\cdot2\cdot10=160$ 个，格式为 NXX 的地区代码有 $8\cdot10\cdot10=800$ 个。类似地，由乘积法则，存在 $8\cdot8\cdot10=640$ 个格式为 NNX 的局代码。乘积法则也表明存在着 $10\cdot10\cdot10\cdot10=10\,000$ 个格式为 XXXX 的话机代码。

因此，再次使用乘积法则，结果在老计划下存在

$$160\cdot640\cdot10\,000=1\,024\,000\,000$$

个不同的北美有效的电话号码。在新计划下存在

$$800\cdot800\cdot10\,000=6\,400\,000\,000$$

个不同的电话号码。■

例 8 执行下面的代码以后 k 的值是什么？

```
k := 0
for i1 := 1 to n1
  for i2 := 1 to n2
      ⋮
    for im := 1 to nm
      k := k+1
```

解 k 的初值是 0。这个嵌套的循环每执行一次，k 就加 1。令 T_i 表示执行第 i 个循环的任务，那么循环执行的次数就是完成任务 T_1，T_2，…，T_m 的方法数。因为对每个整数 i_j，$1\leqslant i_j\leqslant n_j$，第 j 个循环都执行一次，执行任务 $T_j(j=1, 2, \cdots, m)$的方法数就是 n_j。由乘积法则，这个嵌套的循环执行了 $n_1n_2\cdots n_m$ 次。因此 k 最后的值是 $n_1n_2\cdots n_m$。■

例 9 计数有穷集的子集 用乘积法则证明一个有穷集 S 的不同的子集数是 $2^{|S|}$。

解 设 S 是有穷集。按任意的顺序将 S 的元素列成一个表。考虑到在 S 的子集和长为 $|S|$ 的二进制串之间存在着一对一的对应，即如果表的第 i 个元素在这个子集里，则该子集对应的二进制串的第 i 位为 1，否则这位为 0。由乘积法则，存在着 $2^{|S|}$ 个长为 $|S|$ 的二进制串。因此 $|P(S)|=2^{|S|}$。■

乘积法则也常用集合的语言表述如下：如果 A_1，A_2，…，A_m 是有穷集，那么在这些集合的笛卡儿积中的元素数是每个集合的元素数之积。为把这种表述与乘积法则联系起来，注意到在笛卡儿积 $A_1\times A_2\times\cdots\times A_m$ 中选一个元素的任务是通过在 A_1 中选一个元素，A_2 中选一个元素，…，A_m 中选一个元素来完成的。由乘积法则得到

$$|A_1\times A_2\times\cdots\times A_m|=|A_1|\cdot|A_2|\cdot\cdots\cdot|A_m|$$

我们现在引入求和法则。

求和法则 如果完成第一项任务有 n_1 种方式，完成第二项任务有 n_2 种方式，并且这些任务不能同时完成，那么完成第一或第二项任务有 n_1+n_2 种方式。

例 10 说明怎样使用求和法则。

例 10 假定要选一位数学学院的教师或数学专业的学生作为校委会的代表。如果有 37 位数学学院的教师和 83 位数学专业的学生，那么这个代表有多少种不同的选择？

解 完成第一项任务，选一位数学学院的教师，可以有 37 种方式。完成第二项任务，选一位数学专业的学生，有 83 种方式。根据求和法则，结果有 $37+83=120$ 种可能的方式来挑选这个代表。■

我们可以把求和法则推广到多于两项任务的情况。假定任务 T_1，T_2，…，T_m 分别有 n_1，n_2，…，n_m 种完成的方式，并且任何两项任务都不能同时做，那么完成其中一项任务的方式数

是 $n_1+n_2+\cdots+n_m$。正如例 11 和例 12 所示，这个推广的求和法则在计数问题中常常用到。这个求和法则可以从两个集合的求和法则使用数学归纳法加以证明(见本节后的练习 55)。

例 11　一个学生可以从三个表中的一个表选择一个计算机课题。这三个表分别包含 23、15 和 19 个可能的课题。那么课题的选择可能有多少种?

解　这个学生有 23 种方式从第一个表中选择课题，有 15 种方式从第二个表中选择课题，有 19 种方式从第三个表中选择课题。因此，共有 23+15+19=57 种选择课题的方式。■

例 12　在下面的代码执行后 k 的值是什么?

```
k :=0
for i1 :=1 to n1
    k :=k+1
for i2 :=1 to n2
    k :=k+1
      ⋮
for im :=1 to nm
    k :=k+1
```

解　k 的初值是 0。这个代码块由 m 个不同的循环构成。循环中的每次执行 k 都要加 1。令 T_i 是执行第 i 个循环的任务。因为第 i 个循环被执行 n_i 次，所以任务 T_i 可以用 n_i 种方式完成。由于任何两个任务不能同时执行，求和法则证明 k 的最后值，即完成任务 $T_i(i=1,2,\cdots,m)$ 的方式数是 $n_1+n_2+\cdots+n_m$。■

求和法则可以用集合的语言表述如下：如果 A_1，A_2，…，A_m 是不交的集合，那么在其并集中的元素数是每个集合的元素数之和。为把这种表述与求和法则联系起来，令 T_i 是从 $A_i(i=1,2,\cdots,m)$中选取一个元素的任务。有 $|A_i|$ 种方式做 T_i。由于任何两个任务不可能同时做，根据求和法则，从其中某个集合选择一个元素的方式数，即在并集中的元素数是

$$|A_1\cup A_2\cup\cdots\cup A_m|=|A_1|+|A_2|+\cdots+|A_m|$$

这个等式仅适用于问题中的集合是不相交的情况。当这些集合含有公共元素时，情况要复杂得多。本节的后面将对这种情况进行简要的讨论，更深入的讨论放在第 6 章。

4.1.3　比较复杂的计数问题

许多计数问题不能仅仅使用求和法则或者乘积法则来求解。但是，许多复杂的计数问题可以使用这两个法则来求解。

Extra Examples **例 13**　在计算机语言 BASIC 的某个版本中，变量的名字是含有一个或两个字符的符号串，其中的大写和小写字母是不加区分的(一个字符或者取自 26 个英文字母，或者取自 10 个数字)。此外，变量名必须以字母作为开始，并且必须和由两个字符构成的用于程序设计的 5 个保留字相区别。在 BASIC 的这个版本中有多少个不同的变量名?

解　令 V 等于在这个 BASIC 版本中的不同的变量名个数，V_1 是单字符的变量名个数，V_2 是两个字符的变量名个数。那么由求和法则，$V=V_1+V_2$。由于单字符变量名必须是字母，故 $V_1=26$。又根据乘积法则存在 $26\cdot 36$ 个以字母打头以字母数字结尾的 2 位字符串。但是其中 5 个不包含在内，因此 $V_2=26\cdot 36-5=931$。从而在这个 BASIC 版本中存在 $V=V_1+V_2=26+931=957$ 个不同的变量名。■

例 14　计算机系统的每个用户有一个 6～8 个字符构成的密码，其中每个字符是一个大写字母或者数字，且每个密码必须至少包含一个数字。有多少可能的密码?

解 设 P 是可能的密码总数，且 P_6、P_7、P_8 分别表示 6、7 或 8 位的可能的密码数。由求和法则，$P=P_6+P_7+P_8$。我们现在找 P_6、P_7 和 P_8。直接找 P_6 是困难的。而找 6 个大写字母和数字构成的字符串数是容易的，其中包含那些没有数字的串在内，然后从中减去没有数字的串数就得到 P_6。由乘积法则，6 个字符的串数是 36^6，而没有数字的字符串数是 26^6。因此，

$$P_6 = 36^6 - 26^6 = 2\,176\,782\,336 - 308\,915\,776 = 1\,867\,866\,560$$

类似地得到

$$P_7 = 36^7 - 26^7 = 78\,364\,164\,096 - 8\,031\,810\,176 = 70\,332\,353\,920$$

和

$$P_8 = 36^8 - 26^8 = 2\,821\,109\,907\,456 - 208\,827\,064\,576 = 2\,612\,282\,842\,880$$

从而，

$$P = P_6 + P_7 + P_8 = 2\,684\,483\,063\,360$$





例 15 计数因特网网址 在由计算机的物理网络互连而构成的因特网中，每台计算机（或者更精确地说是计算机的每个网络连接）被分配一个因特网地址。目前正在使用的因特网协议版本 4(IPv4)中，一个地址是一个 32 位的二进制串。它以网络标识(netid)开始，后面跟随着主机标识(hostid)，该标识把一个计算机认定为某个指定网络的成员。

根据网络标识和主机标识位数的不同，使用 3 种地址形式。用于最大网络的 *A 类地址*，由 0 后跟 7 位的网络标识和 24 位的主机标识构成。用于中等规模网络的 *B 类地址*，由 10 后跟 14 位的网络标识和 16 位的主机标识构成。用于最小网络的 *C 类地址*，由 110 后跟 21 位的网络标识和 8 位的主机标识构成。由于特定用途，对地址有着某些限制：1111111 在 A 类网络的网络标识中是无效的，全 0 和全 1 组成的主机标识对任何网络都是无效的。因特网上的一台计算机有一个 A 类、B 类或 C 类地址。（除了 A 类、B 类和 C 类地址外，还有 D 类地址和 E 类地址。D 类地址在多台计算机同时编址时用于多路广播，它由 1110 后跟 28 位组成。E 类地址保留为将来应用，由 11110 后跟 27 位组成。D 和 E 类地址不会分配给因特网中的一台计算机作为 IP 地址。）图 4-1 显示了 IPv4 的编址。（A 类和 B 类网络标识的数量限制已经使得 IPv4 编址不够用了；将代替 IPv4 的 IPv6 使用 128 位地址来解决这个问题。）

对因特网上的计算机有多少不同的有效 IPv4 地址？

位数	0	1	2	3	4	8	16	24	31
A类	0	网络标识				主机标识			
B类	1	0	网络标识				主机标识		
C类	1	1	0	网络标识				主机标识	
D类	1	1	1	0	多路广播地址				
E类	1	1	1	1	0	地址			

图 4-1 因特网地址(IPv4)

解 令 x 是因特网上计算机的有效地址数，x_A、x_B 和 x_C 分别表示 A 类、B 类和 C 类的有效地址数。由求和法则，$x=x_A+x_B+x_C$。

为了找到 x_A，由于 1111111 是无效的，故存在 $2^7-1=127$ 个 A 类的网络标识。对于每个网络标识，存在 $2^{24}-2=16\,777\,214$ 个主机标识，这是由于全 0 和全 1 组成的主机标识是无效的。因此，$x_A=127\cdot 16\,777\,214=2\,130\,706\,178$。

为了找到 x_B 和 x_C，首先注意到存在 $2^{14}=16\,384$ 个 B 类网络标识和 $2^{21}=2\,097\,152$ 个 C 类

网络标识。对每个 B 类网络标识存在着 $2^{16}-2=65\ 534$ 个主机标识，而对每个 C 类网络标识存在着 $2^8-2=254$ 个主机标识，这也是考虑到全 0 和全 1 组成的主机标识是无效的。因而，$x_B=1\ 073\ 709\ 056$，$x_C=532\ 676\ 608$。

我们可以断言 IPv4 有效地址的总数是 $x=x_A+x_B+x_C=2\ 130\ 706\ 178+1\ 073\ 709\ 056+532\ 676\ 608=3\ 737\ 091\ 842$。■

4.1.4 容斥原理

当同时做两个任务时，我们不能使用求和法则来计数完成其中一个任务的方式。把对每个任务的方式数加起来将导致计数结果的增大，因为同时完成两个任务的那些方式被计数了两次。为了正确地计数完成其中一个任务的方式，我们先把完成每个任务的方式数加起来，然后再减去同时完成两个任务的方式数。这个技术叫做*容斥原理*。例 16 显示了怎样用这个原理来求解计数问题。

例 16　以 1 开始或者以 00 结束的 8 位二进制符号串有多少个？

解　第一个任务，构造以 1 开始的 8 位二进制字符串，完成它有 $2^7=128$ 种方式，这是由乘积法则得到的。因为第一位只有一种选择方式，而其他 7 位中的每位有两种选择方式。

第二个任务，构造以 00 结束的 8 位二进制字符串，完成它有 $2^6=64$ 种方式，这也是由乘积法则得到的。因为前 6 位的每位有两种选择方式，而最后两位只有一种选择方式。

同时完成两个任务，构造以 1 开始以 00 结束的 8 位二进制符号串，完成它有 $2^5=32$ 种方式。这里也使用了乘积法则，因为第一位只有一种选法，从第二位到第六位每位可以有两种选法，最后两位也只有一种选法。因而，以 1 开始或者以 00 结束的 8 位二进制符号串个数，即完成第一或第二个任务的方式数，等于 $128+64-32=160$。■

我们可以用集合的语言表述这一计数原理。令 A_1 和 A_2 是集合，T_1 是从 A_1 选择一个元素的任务，T_2 是从 A_2 选择一个元素的任务。完成 T_1 有 $|A_1|$ 种方式，完成 T_2 有 $|A_2|$ 种方式。完成 T_1 或 T_2 的方式数是完成 T_1 的方式数与完成 T_2 的方式数之和减去同时完成 T_1、T_2 两个任务的方式数。因为存在 $|A_1\cup A_2|$ 种方式完成 T_1 或 T_2，$|A_1\cap A_2|$ 种方式完成 T_1 和 T_2，我们有

$$|A_1\cup A_2|=|A_1|+|A_2|-|A_1\cap A_2|$$

这就是在 1.7 节给出的计数两个集合并集中元素的公式。

容斥原理可以推广来求完成 n 个不同任务中的一个任务的方式数，换句话说，就是找 n 个集合的并集中的元素数，其中 n 是正整数。我们将在第 6 章研究容斥原理和它的某些广泛应用。

4.1.5 树图

可以使用树图求解计数问题。一棵树由根、从根出发的许多分支以及可能从其他分支端点出发的新的分支构成。（我们将在第 9 章详细地研究树）。为了在计数中使用树，我们用一个分支表示每个可能的选择，用树叶表示可能的结果。这些树叶是某些分支的端点，从这些端点不再进一步分支。

注意，当用树图求解计数问题时，为到达一片树叶所作的选择个数可能是不同的（作为例子，见例 18）。

例 17　有多少不含连续两个 1 的 4 位二进制串？

解　图 4-2 的树图给出了所有不含连续两个 1 的 4 位二进制串。我们看出存在 8 个不含连续两个 1 的 4 位二进制串。■

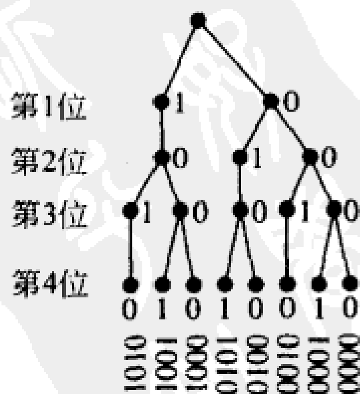


图 4-2　不含连续两个 1 的 4 位二进制串

例 18　在两个队（队 1 和队 2）之间的决赛至多由 5 次比赛构成。先胜 3 次的队赢得决赛。决赛可能出现多少种不同的方式？

解 在图 4-3 的树图中，以每次比赛的得胜者给出了决赛可能进行的所有方式。我们看到有 20 种不同的决赛的方式。■

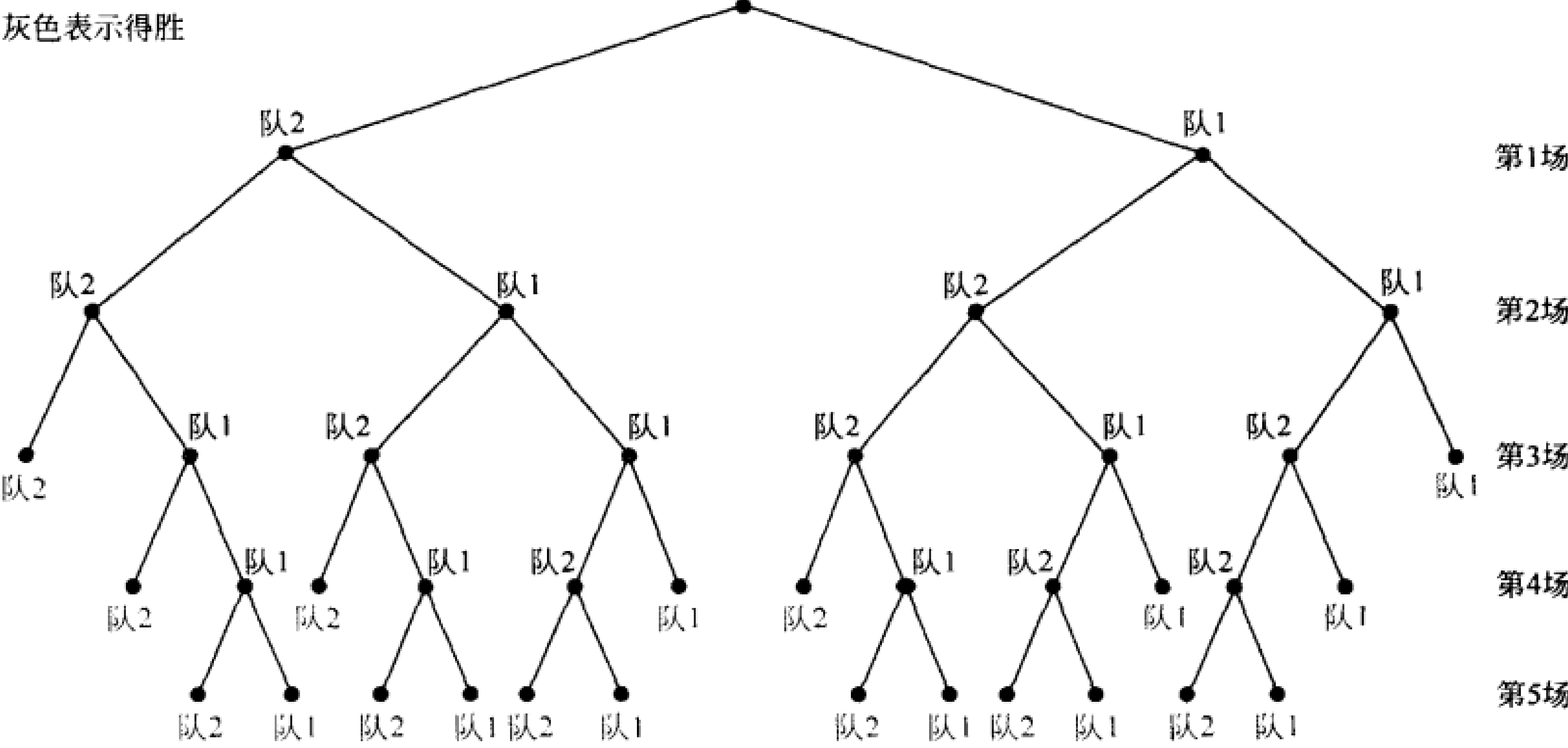


图 4-3 5 次决赛胜 3 次

例 19 假设"我爱新泽西"T 恤衫有 5 种不同的规格：S、M、L、XL 和 XXL。又知道 XL 规格只有三种颜色，红色、绿色和黑色，XXL 规格只有绿色和黑色。除此之外，其他规格有四种颜色，白色、红色、绿色和黑色。如果每种规格和颜色的 T 恤衫至少一件，一个纪念品商店必须库存多少件不同的 T 恤衫？

解 图 4-4 的树图给出了所有规格和颜色的配对。从中可知这个纪念品商店老板必须库存 17 件不同的 T 恤衫。■

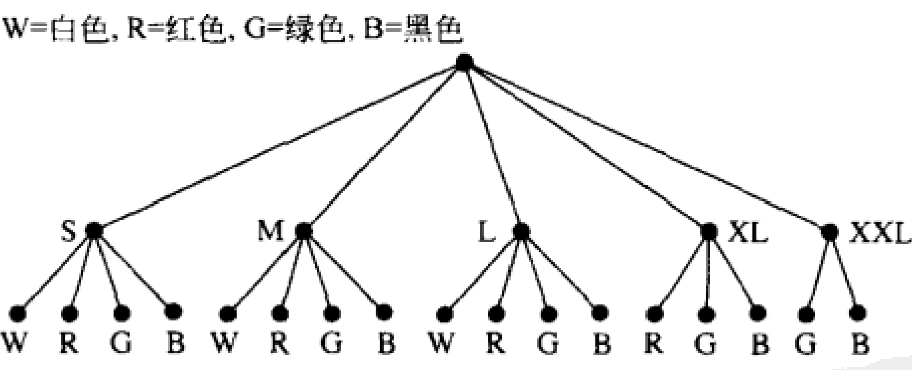


图 4-4 计数不同的 T 恤衫

练习

1. 一个学院有 18 个数学专业和 325 个计算机科学专业的学生。
   a)选两个代表，使得一个是数学专业的且另一个是计算机科学专业的，有多少种方式？
   b)选一个数学专业或计算机科学专业的代表又有多少种方式？
2. 一个办公大楼有 27 层，每层有 37 个办公室，那么在这个大楼里有多少个办公室？
3. 一次多项选择考试包含 10 个问题。每个问题有 4 个可能的答案。
   a)在这次考试中如果每个问题都要回答，一个学生回答这些问题可能有多少种方式？
   b)在这次考试中如果允许某些答案空缺，一个学生回答这些问题可能有多少种方式？

4. 某种商标的衬衫有 12 种颜色，有男式和女式 2 种样式，每种样式有 3 种大小型号。问这些衬衫有多少种不同的类型？
5. 从纽约到丹佛有 6 条不同的航线，而从丹佛到旧金山有 7 条。如果选一个到丹佛的航班，接着选一个到旧金山的航班，那么从纽约经丹佛到旧金山的旅行有多少种不同的可能性？
6. 从波士顿到底特律有 4 条汽车主干线，而从底特律到洛杉矶有 6 条。那么从波士顿经底特律到洛杉矶的汽车主干线有多少条？
7. 如果用 3 个字母作为姓名的缩写，人们可以有多少种不同的选择？
8. 如果这 3 个字母不允许重复，人们可以有多少种不同的选择？
9. 如果这 3 个字母以 A 开始，人们又可以有多少种不同的选择？
10. 8 位二进制串有多少个？
11. 首尾都是 1 的 10 位二进制串有多少个？
12. 位数不超过 6 的二进制串有多少个？
13. 位数不超过 n 且全由 1 组成的二进制串有多少个？这里的 n 是正整数。
14. 首尾都是 1 的 n 位二进制串有多少个？这里的 n 是正整数。
15. 位数不超过 4 且由小写字母构成的串有多少个？
16. 由 4 个小写字母构成且含有字母 x 的串有多少个？
17. 由 5 个 ASCII 码构成且至少(在符号位)包含一个@字符的串有多少个？(注：有 128 个不同的 ASCII 码)
18. 有多少个小于 1000 的正整数
    a)被 7 整除？
    b)被 7 整除但不被 11 整除？
    c)同时被 7 和 11 整除？
    d)被 7 或 11 整除？
    e)恰好被 7 或 11 中的一个数整除？
    f)既不被 7 整除，也不被 11 整除？
    g)含有不同的数字？
    h)含有不同的数字且是偶数？
19. 在 100 到 999 之间包含多少个正整数
    a)被 7 整除？
    b)是奇数？
    c)有相同的 3 个十进制数字？
    d)不被 4 整除？
    e)被 3 或 4 整除？
    f)不被 3 也不被 4 整除？
    g)被 3 整除但不被 4 整除？
    h)被 3 和 4 整除？
20. 在 1000 到 9999 之间包含多少个正整数
    a)被 9 整除？
    b)是偶数？
    c)有不同的十进制数字？
    d)不被 3 整除？
    e)被 5 或 7 整除？
    f)不被 5 也不被 7 整除？
    g)被 5 整除但不被 7 整除？
    h)被 5 和 7 整除？
21. 有多少个串含有 3 个十进制数字且
    a)同一数字不能出现 3 次？
    b)以奇数字开始？
    c)恰有 2 个数字是 4？
22. 有多少个串含有 4 个十进制数字且
    a)同一数字不出现两次？
    b)以偶数字结束？
    c)恰有 3 个数字是 9？
23. 一个委员会由 50 个州构成，每个州可从州长或两个参议员中选一个参加，有多少种不同的方式？
24. 用 3 个数字后跟 3 个字母或者 3 个字母后跟 3 个数字可构成多少种车牌？
25. 用 2 个字母后跟 4 个数字或者 2 个数字后跟 4 个字母可构成多少种车牌？
26. 用 3 个字母后跟 3 个数字或者 4 个字母后跟 2 个数字可构成多少种车牌？
27. 用 2 个或 3 个字母后跟 2 个或 3 个数字可构成多少种车牌？
28. 由 8 个英语字母可构成多少个串？
    a)如果字母可以重复
    b)如果字母不能重复

c)如果字母可以重复且以 X 开始
d)如果字母不能重复且以 X 开始
e)如果字母可以重复且以 X 开始和结束
f)如果字母可以重复且以 BO(按此次序)开始
g)如果字母可以重复且以 BO(按此次序)开始和结束
h)如果字母可以重复且以 BO(按此次序)开始或结束

29. 由 8 个英语字母可构成多少个串?
a)如果字母可以重复且不包含元音字母
b)如果字母不能重复且不包含元音字母
c)如果字母可以重复且以元音字母开始
d)如果字母不能重复且以元音字母开始
e)如果字母可以重复且包含至少一个元音字母
f)如果字母可以重复且包含恰好一个元音字母
g)如果字母可以重复且以 X 开始并至少包含一个元音字母
h)如果字母可以重复且以 X 开始和结束并至少包含一个元音字母

30. 从 10 元素集合到含有下述元素数的集合有多少个不同的函数?
a)2 b)3 c)4 d)5

31. 从 5 元素集合到含有下述元素数的集合有多少一对一的函数?
a)4 b)5 c)6 d)7

32. 从集合{1, 2, …, n}到集合{0, 1}有多少个函数? 这里的 n 是正整数。

33. 从集合{1, 2, …, n}到集合{0, 1}有多少个满足下列条件的函数? 这里的 n 是正整数。
a)是一对一的 b)对 1 和 n 赋值为 0
c)对恰好一个小于 n 的正整数赋值为 1

34. 从 5 元素集合到含有下述元素数的集合有多少个部分函数(见 1.8 节的练习)?
a)1 b)2 c)5 d)9

35. 从 m 元素集合到 n 元素集合有多少个部分函数(见 1.8 节的练习)? 这里的 m 和 n 是正整数。

36. 100 个元素的集合有多少个子集的元素数多于 1?

37. 如果一个字符串反转后所得结果与原来的字符串一样，就称它是一个回文。有多少个长为 n 的二进制串是回文?

38. 在一个婚礼上摄影师从 10 个人中安排 6 个人在一排拍照，其中新娘和新郎也在这 10 个人中，如果满足下述条件，有多少种安排方式?
a)新娘必须在照片中 b)新娘和新郎必须都在照片中
c)新娘和新郎恰好有一个在照片中

39. 在一个婚礼上摄影师安排 6 个人在一排拍照，包含新娘和新郎在内，如果满足下述条件，有多少种安排方式?
a)新娘必须在新郎旁边 b)新娘不在新郎旁边
c)新娘在新郎左边的某个位置

40. 有多少个 7 位二进制串以 2 个 0 开始或以 3 个 1 结束?

41. 有多少个 10 位二进制串以 3 个 0 开始或以 2 个 0 结束?

*42. 有多少个 10 位二进制串包含 5 个连续的 0 或者 5 个连续的 1?

**43. 有多少个 8 位二进制串包含 3 个连续的 0 或者 4 个连续的 1?

44. 离散数学班的每个学生都是计算机科学或数学专业的，或者是同时修这两个专业的。如果有 38 个是计算机科学专业的(包含同时修两个专业的)，23 个是数学专业的(包含同时修两个专业的)，7 个同时修两个专业的，那么这个班有多少个学生?

45. 有多少个不超过 100 的正整数能被 4 或 6 整除?

46. 在 C 程序设计语言中的变量名是一个字符串，可以包含大写字母、小写字母、数字或下划线。此外，字符串的第一个字符必须是字母(大写或小写字母)或下划线。如果一个变量名由它的前 8 个字符确定，那么在 C 语言中可以命名多少个不同的变量?（注：变量名包含的字符数可以少于 8 个）。
47. 假定在将来的某个时间世界上的每部电话将被分配一个号码，这个号码包含一个 1 到 3 位数字的形如 X、XX 或 XXX 的国家代码，后面跟随着一个 10 位数字的形如 NXX-NXX-XXXX 的电话号码(如例 7 所描述的)。在这个编码计划中，全世界将有多少个不同的有效电话号码?
48. 使用树图找出不含 3 个连续 0 的 4 位二进制串的个数。
49. 有多少种不同的方式排列字母 a、b、c 和 d，使得 b 不紧跟在 a 的后边?
50. 使用树图找出世界职业棒球大赛可能出现的方式数，其中 7 场中先胜 4 场的队赢得这个比赛。
51. 使用树图确定{3，7，9，11，24}的子集数，使得子集中的元素之和小于 28。
52. a)假设一个商店出售 6 种不同的软饮料：可乐、姜汁茶、橙汁、乐啤露、柠檬茶和奶油苏打。所有品种的瓶装饮料都有 12 盎司规格的。除了柠檬茶外其他品种都有 20 盎司规格的。只有可乐和姜汁茶有 32 盎司规格。除了柠檬茶和奶油苏打以外其他品种都有 64 盎司规格的。如果这个商店要具有所有品种和规格的饮料，使用树图确定它必须库存多少瓶不同的饮料?

    b)使用计数原理回答(a)中的问题。
53. a)假设运动鞋的流行式样对男女都适用。女鞋的大小号码是 6，7，8，9，男鞋的大小号码是 8，9，10，11 和 12。男鞋有白色和黑色，而女鞋是白色、红色和黑色。如果一个商店各种大小和颜色的男、女运动鞋必须至少存一双，用树图确定所需的鞋的数目。

    b)使用计数原理回答(a)中的问题。

*54. 使用乘积法则证明对于 n 个变量的命题存在 2^{2^n} 个不同的真值表。

55. 从两个任务的求和法则使用数学归纳法证明关于 m 个任务的求和法则。
56. 从两个任务的乘积法则使用数学归纳法证明关于 m 个任务的乘积法则。
57. 具有 n 条边的凸多边形有多少条对角线?（如果在多边形内或边界的每两个顶点的连线完全在这个集合之内，则称为凸多边形)。
58. 因特网中的数据以*数据报*传输，数据报是由二进制位的数据块构成的。每个数据报包含有头信息和数据区。头信息最多被分成 14 个不同的域(详细说明许多事项，包括发送和接收地址)，数据区包含被传输的实际数据。14 个头信息域中有一个*头长度域*(表示为 HLEN)，根据协议规定是 4 位，它说明了以 32 位为一个数据块的头信息的长度。例如，如果 HLEN＝0110，那么头信息由 6 个 32 位的数据块构成。14 个头信息域中的另一个域是 16 位的*总长域*(表示为 TOTAL LENGTH)，它说明了以位为单位的整个数据报包含头信息和数据区在内的总长度。数据区的长度是数据报的总长度减去头的长度。

    a)TOTAL LENGTH 的最大值(16 位长)确定了因特网数据报以字节(8 位的数据块)为单位的最大总长度。这个值是多少?

    b)HLEN 的最大值(4 位长)确定了头信息以 32 位数据块为单位的最大总长度，这个值是多少?以字节为单位的最大的头信息的总长度是多少?

    c)最小的(最通常的)头长度是 20 字节。因特网数据报的数据区以字节为单位的最大总长度是多少?

    d)如果头长度是 20 个字节并且总长度是尽可能地长，那么在数据区可以传输多少个不同的字节串?

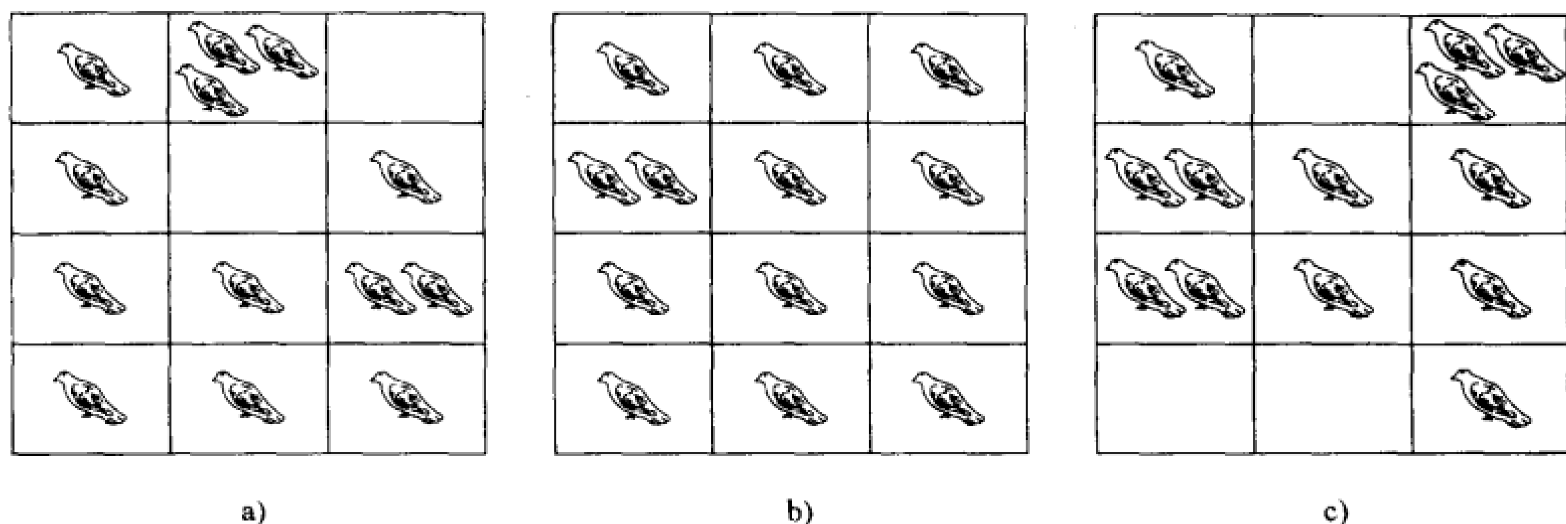
4.2 鸽巢原理

4.2.1 引言



假定一群鸽子飞入一组鸽巢安歇。*鸽巢原理*表明如果鸽子数比鸽巢数多，那么一定有一个鸽巢里至少有 2 只鸽子(见图 4-5)。当然，这个原理除了鸽子和鸽巢外也可以用于其他对象。

定理 1　*鸽巢原理*　如果 $k+1$ 个或更多的物体放入 k 个盒子，那么至少有一个盒子包含了 2 个或更多的物体。



a) b) c)

图 4-5 鸽子比鸽巢多

证 假定 k 个盒子中没有一个盒子包含的物体多于 1 个，那么物体总数至多是 k，这与至少有 $k+1$ 个物体矛盾。 □

鸽巢原理也叫做*狄利克雷抽屉原理*，以 19 世纪的德国数学家狄利克雷[⊖]命名，他经常在工作中使用这个原理。下面的例子说明了怎样使用鸽巢原理。

例 1 在一组 367 个人中一定至少有 2 个人有相同的生日，这是由于只有 366 个可能的生日。 ■

例 2 在 27 个英文单词中一定至少有 2 个单词以同一个字母开始，因为英文字母表中只有 26 个字母。 ■

例 3 如果考试给分是从 0 到 100，班上必须有多少个学生才能保证在这次期末考试中至少有 2 个学生得到相同的分数？

解 期末考试有 101 个分数。鸽巢原理证明在 102 个学生中一定至少有 2 个学生具有相同的分数。 ■

鸽巢原理在许多证明中都是有用的工具，有些证明结果是令人意外的，正如例 4 所给出的。

例 4 证明对每个整数 n，存在一个数是 n 的倍数，且在它的十进制表示中只出现 0 和 1。

Extra Examples **解** 令 n 是正整数。考虑 n 个整数 1，11，111，…，11…1(在这个数表中，最后一个整数的十进制表示中具有 $n+1$ 个 1)。注意到当一个整数被 n 整除时存在 n 个可能的余数。因为这个数表中有 $n+1$ 个整数，由鸽巢原理必有两个整数在除以 n 时有相同的余数。这两个整数之差的十进制表示中只含有 0 和 1，且它能被 n 整除。 ■

4.2.2 广义鸽巢原理

鸽巢原理指出当物体比盒子多时一定至少有 2 个物体在同一个盒子里。但是当物体数超过盒子数的倍数时可以得出更多的结果。例如，在任意 21 个十进制数字中一定有 3 个是相同的。这是由于 21 个物体被分配到 10 个盒子里，那么某个盒子的物体一定多于 2 个。

定理 2 *广义鸽巢原理* 如果 N 个物体放入 k 个盒子，那么至少有一个盒子包含了至少

⊖ **G. L. 狄利克雷(G. Lejenue Dirichlet，1805—1859)** 狄利克雷诞生在德国科隆附近的一个法国家庭。他在巴黎大学学习并且在布莱斯劳大学和柏林大学工作。1855 年他在哥廷根大学继高斯的职位。据说狄利克雷是第一个掌握 20 年前高斯的“算学研究”的人。据说他手边总有一份抄件，即使旅行时也不例外。狄利克雷在数论中有许多重要的发现，包括在算术级数 $an+b$ 中存在无限多素数的定理，这里的 a 和 b 互素。他对于 $n=5$ 的情况证明了费马大定理，即方程 $x^5+y^5=z^5$ 不存在非平凡的整数解。狄利克雷也对分析做出了许多贡献。

$\lceil N/k \rceil$个物体。

证 假定没有盒子包含了比$\lceil N/k \rceil-1$多的物体，那么物体总数至多是

$$k\left(\left\lceil \frac{N}{k} \right\rceil - 1\right) < k\left(\left(\frac{N}{k}+1\right)-1\right) = N$$

这里用到不等式$\lceil N/k \rceil < (N/k)+1$。这与存在有总数 N 个物体矛盾。 □

一类普遍的问题是，把一些物体分到 k 个盒子中要使得某个盒子至少含有 r 个物体，求这些物体的最少个数。当有 N 个物体时，广义鸽巢原理告诉我们，只要$\lceil N/k \rceil \geq r$一定有 r 个物体在同一个盒子里。满足 $N/k > r-1$ 的最小正整数，即 $N=k(r-1)+1$，是满足不等式$\lceil N/k \rceil \geq r$的最小正整数。还可能有更小的 N 值吗？答案是没有。原因是如果我们有 $k(r-1)$个物体，我们可以在 k 个盒子中的每个盒子放 $r-1$ 个物体，因此没有一个盒子至少有 r 个物体。

当思考这种问题时下面的想法是有用的，就是在不断地放物体时怎样避免一个盒子至少有 r 个物体出现。为避免把第 r 个物体加到任何一个盒子里，每个盒子最终将以具有 $r-1$ 个物体结束。如果不允许放第 r 个物体到一个盒子里，就没有办法增加下一个物体。

例 5 至例 8 说明了怎样使用广义鸽巢原理。

例 5 在 100 个人中至少有$\lceil 100/12 \rceil = 9$ 个人生在同一个月。 ■

例 6 如果有 5 个可能的成绩 A、B、C、D 和 F，那么在一个离散数学班里最少要多少个学生才能保证至少 6 个学生得到相同的分数？

解 为保证至少 6 个学生得到相同的分数，所需的最少学生数是使得$\lceil N/5 \rceil = 6$ 的最小整数 N。这样的最小整数是 $N=5\cdot 5+1=26$。如果只有 25 个学生，可能是每 5 个学生得到同样的分数，而没有 6 个学生得到同样的分数。于是，26 是保证至少 6 个学生得到相同的分数所需的最少学生数。 ■

例 7 a)从一副标准的 52 张牌中必须选多少张牌才能保证选出的牌中至少有 3 张是同样的花色？

b)必须选多少张牌才能保证选出的牌中至少有 3 张是红心？

解 a)假设存在四个盒子保存四种花色的牌，选中的牌放在同种花色的盒子里。使用广义鸽巢原理，如果选了 N 张牌，那么至少有一个盒子含有至少$\lceil N/4 \rceil$张牌。因此如果$\lceil N/4 \rceil \geq 3$，我们知道至少选了 3 张同种花色的牌。使得$\lceil N/4 \rceil \geq 3$ 的最小的整数 N 是 $N=2\cdot 4+1=9$，所以 9 张牌就足够了。注意到，如果选 8 张牌，可能每种花色 2 张牌。因而必须选 9 张牌才能保证选出的牌中至少 3 张是同样的花色。想到这一点的一个好方法就是注意到在选了 8 张牌以后没有办法避免出现 3 张同样花色的牌。

b)我们不用广义鸽巢原理回答这个问题，因为我们要保证存在 3 张红心而不仅仅是 3 张同样花色的牌。注意到在最坏情况下，选一张红心以前可能选了所有的黑桃、方块、梅花，总共 39 张牌，下面选的 3 张牌将都是红心。因此为得到 3 张红心，可能需要选 42 张牌。 ■

例 8 为保证一个州的 2500 万个电话有不同的 10 位电话号码，所需地区代码的最小数是多少？（假定电话号码是 NXX-NXX-XXXX 形式，其中前 3 位是地区代码，N 表示从 2 到 9 包含的十进制数字，X 表示任何十进制数字）。

解 有 800 万个形如 NXX-XXXX 的不同的电话号码（如 4.1 节的例 7 所示）。因此，由广义鸽巢原理，在 2500 万个电话号码中，一定至少有$\lceil 25\,000\,000/8\,000\,000 \rceil$个同样的电话号码。因而至少需要 4 个地区代码来保证所有的 10 位号码是不同的。 ■

例 9 尽管没有用到广义鸽巢原理，也用到了类似的原理。

例 9 假设计算机科学实验室有 15 台工作站和 10 台服务器。可以用一条电缆直接把工作站连到服务器。同一时刻只有一条到服务器的直接连接是有效的。我们想保证在任何时刻任何一组不超过 10 个工作站可以通过直接连接同时访问不同的服务器。尽管我们可以通过将每台工作站直接连接到每台服务器(使用 150 条连线)来做到这一点，达到这个目标所需要的最少直接连线的数目是多少?

解 将工作站标记为 W_1，W_2，…，W_{15}，服务器标记为 S_1，S_2，…，S_{10}。假设对于 $k=1, 2, \cdots, 10$，我们连接 W_k 到 S_k，并且 W_{11}，W_{12}，W_{13}，W_{14} 和 W_{15} 中的每个工作站都连接到所有的 10 个服务器。总共 60 条直接连线。很清楚，在任何时刻任何一组不超过 10 个工作站可以通过直接连接同时访问不同的服务器。为看到这一点只要注意到下述事实：如果这个组包含工作站 $W_j(1\leqslant j\leqslant 10)$，那么 W_j 可以访问服务器 S_j。对于组里的每个工作站 $W_k(k\geqslant 11)$，一定存在不在组里的工作站 $W_j(1\leqslant j\leqslant 10)$ 与之对应，因此 W_k 可以访问服务器 S_j(这是由于存在多少个不在组里的工作站 W_j，$1\leqslant j\leqslant 10$，至少存在同样多的服务器 S_j 可以被其他工作站访问)。

现在假设在工作站和服务器之间直接连线少于 60 条。那么某个服务器将至多连接 $\lfloor 59/10\rfloor=5$ 工作站。(如果所有的服务器连接到至少 6 个工作站，那么将存在至少 $6\cdot 10=60$ 条直接连线。)这意味着剩下的 9 个服务器对于其他 10 个工作站同时访问不同的服务器就不够用了。因此，至少需要 60 条直接连线，从而得到答案是 60。■

4.2.3 巧妙使用鸽巢原理

在鸽巢原理的许多有趣应用中，必须用某种巧妙的方式选择放入盒子的物体。下面将描述这样的一些应用。

例 10 在 30 天的一个月里，某棒球队一天至少打一场比赛，但至多打 45 场。证明一定有连续的若干天内这个队恰好打了 14 场。

解 令 a_j 是在这个月的第 j 天或第 j 天之前所打的场数，则 a_1，a_2，…，a_{30} 是不同正整数的一个递增序列，其中 $1\leqslant a_j\leqslant 45$。从而 a_1+14，a_2+14，…，$a_{30}+14$ 也是不同正整数的一个递增序列，其中 $15\leqslant a_j+14\leqslant 59$。

60 个正整数 a_1，a_2，…，a_{30}，a_1+14，a_2+14，…，$a_{30}+14$ 全都小于等于 59。因此，由鸽巢原理有两个正整数相等。因为整数 a_j，$j=1, 2, \cdots, 30$ 都不相同，并且 a_j+14，$j=1, 2, \cdots, 30$ 也不相同，一定存在下标 i 和 j 满足 $a_i=a_j+14$。这意味着从第 $j+1$ 天到第 i 天恰好打了 14 场比赛。■

例 11 证明在不超过 $2n$ 的任意 $n+1$ 个正整数中一定存在一个正整数被另一个正整数整除。

解 把 $n+1$ 个整数 a_1，a_2，…，a_n 中的每一个都写成 2 的幂与一个奇数的乘积。换句话说，令 $a_j=2^{k_j}q_j$，$j=1, 2, \cdots, n+1$，其中 k_j 是非负整数 q_j 是奇数。整数 q_1，q_2，…，q_{n+1} 都是小于 $2n$ 的正奇数。因为只存在 n 个小于 $2n$ 的正奇数，由鸽巢原理，q_1，q_2，…，q_{n+1} 中必有两个相等。于是，存在整数 i 和 j 使得 $q_i=q_j$。令 q_i 与 q_j 的公共值是 q，那么 $a_i=2^{k_i}q$，$a_j=2^{k_j}q$。因而，若 $k_i<k_j$，则 a_i 整除 a_j；若 $k_i>k_j$，则 a_j 整除 a_i。■

巧妙地应用鸽巢原理证明了在不同整数的序列中存在着确定长度的递增或递减子序列。在给出这个应用之前先回顾某些定义。假定 a_1，a_2，…，a_N 是实数序列。它的一个*子序列*是形为 a_{i1}，a_{i2}，…，a_{im} 的序列，其中 $1\leqslant i_1<i_2<\cdots<i_m\leqslant N$。因此一个子序列是从初始序列得到的序列，按照原来的顺序选取初始序列的某些项，也许要排除其他的项。如果这个序列的每一项都大于它前面的项，就称为*严格递增的*，如果每一项都小于它前面的项，就称为*严格递减的*。

定理 3 每个由 n^2+1 个不同实数构成的序列都包含一个长为 $n+1$ 的严格递增子序列或严

格递减子序列。

在证明定理 3 之前先给出一个例子。

例 12　序列 8，11，9，1，4，6，12，10，5，7 包含 10 项。注意到 $10=3^2+1$，存在四个长为 4 的递增子序列，即 1，4，6，12；1，4，6，7；1，4，6，10 和 1，4，5，7。也存在一个长为 4 的递减子序列，即 11，9，6，5。■

现在给出定理的证明。

证　令 a_1，a_2，…，a_{n^2+1} 是 n^2+1 个不同实数的序列。和序列中的每一项 a_k 联系着一个有序对，即(i_k, d_k)，其中 i_k 是从 a_k 开始的最长的递增子序列的长度，且 d_k 是从 a_k 开始的最长的递减子序列的长度。

假定没有长为 $n+1$ 的递增或递减子序列。那么 i_k 和 d_k 都是小于或等于 n 的正整数，$k=1$，2，…，n^2+1。因此，由乘积法则关于(i_k, d_k)存在 n^2 个可能的有序对。根据鸽巢原理，n^2+1 个有序对中必有两个相等。换句话说，存在项 a_s 和 a_t，$s<t$，使得 $i_s=i_t$ 和 $d_s=d_t$。我们将证明这是不可能的。由于序列的项是不同的，不是 $a_s<a_t$ 就是 $a_s>a_t$。如果 $a_s<a_t$，那么由于 $i_s=i_t$，那么把 a_s 加到从 a_t 开始的递增子序列前面就构造出一个从 a_s 开始的长为 i_s+1 的递增子序列。从而产生矛盾。类似地，如果 $a_s>a_t$，可以证明 d_s 一定大于 d_t，从而也产生矛盾。□



最后的例子说明了怎样把广义鸽巢原理用于组合学的重要部分**拉姆齐理论**，它是以英国数学家拉姆齐[㊀]而命名的。一般地说，拉姆赛理论可用于处理集合元素的子集分配问题。

例 13　假定一组有 6 个人，任两个人或者是朋友或者是敌人。证明在这组人中或存在 3 个人彼此都是朋友，或存在 3 个人彼此都是敌人。

解　令 A 是 6 个人中的一人，组里其他 5 个人中至少有 3 个人是 A 的朋友，或至少有 3 个人是 A 的敌人。这可从广义鸽巢原理得出，因为当 5 个物体分成两个集合时，其中的一个集合至少有$\lceil 5/2\rceil=3$ 个元素。若是前一种情况，假定 B、C 和 D 是 A 的朋友。如果这 3 个人中有 2 个人也是朋友，那么这 2 个人和 A 构成彼此是朋友的 3 人组。否则，B、C 和 D 构成彼此为敌人的 3 人组。对于后一种情况的证明，当 A 存在 3 个或更多的敌人时可以用类似的方法处理。■

拉姆齐数 $R(m, n)$(其中 m 和 n 是大于或等于 2 的正整数)表示：假设晚会上每两个人是朋友或者是敌人，那么在一个晚会上使得或者有 m 个人两两都是朋友，或者有 n 个人两两都是敌人所需要的最少人数。例 13 显示 $R(3, 3)\leqslant 6$。在一组 5 个人中，其中每两个人是朋友或是敌人，可能没有 3 个人两两是朋友，也没有 3 个人两两是敌人，因此我们断言 $R(3, 3)=6$(见练习 24)。

可以证明某些关于拉姆齐数的有用的性质，但是对于大多数拉姆齐数找到精确的值是困难的。根据对称性可以证明 $R(m, n)=R(n, m)$(见练习 28)。对于每个正整数 $n\geqslant 2$ 我们也有 $R(2, n)=n$(见练习 27)。只知道 9 个拉姆齐数 $R(m, n)$$(3\leqslant m\leqslant n)$的精确值，其中包括 $R(4, 4)=18$。对许多其他的拉姆齐数只知道界，包括 $R(5, 5)$在内，已知它满足 $43\leqslant R(5, 5)\leqslant 49$。有兴



㊀ **富兰克·波拉姆顿·拉姆齐(Frank Plumpton Ramsey，1903—1930)**　拉姆齐是剑桥马格达林学院校长的儿子，在温彻斯特和特里尼特学院受过教育。1923 年毕业以后，他应聘在剑桥皇家学院工作，并在那里度过余生。拉姆齐对数理逻辑作出了重要的贡献。我们现在所称的拉姆齐理论是由他在“一个形式逻辑问题”的论文中所发表的聪明的组合论辩引起的。拉姆齐也对经济数学理论作出了贡献。他作为在数学基础方面的优秀讲师而受到注意。他死于 26 岁，他的死使得数学界和剑桥大学失去了一个才华横溢的年轻学者。

趣了解更多有关拉姆齐数知识的读者可以参考[MiRo91]或[GrRoSp90]。

练习

1. 假定周末不排课，证明在任一组 6 门课中一定有 2 门课安排在同一天上课。
2. 如果一个班有 30 个学生，证明至少 2 个学生的姓以同一个字母开头。
3. 抽屉里有一打棕色的短袜和一打黑色的短袜，全都没有配好对。一个人在黑暗中随机取出一些袜子。
   a)必须取多少只袜子才能保证至少有 2 只袜子是同色的？
   b)必须取多少只袜子才能保证至少有 2 只袜子是黑色的？
4. 一个碗里有 10 个红球和 10 个蓝球。一个女士不看着球而随机地选取。
   a)她必须选多少个球才能保证至少有 3 个球是同色的？
   b)她必须选多少个球才能保证至少有 3 个球是蓝色的？
5. 证明在任意 5 个整数中(不一定是连续的)有 2 个整数被 4 除的余数相等。
6. 设 d 是正整数，证明在任意一组 $d+1$ 个整数中(不一定是连续的)有 2 个整数被 d 除的余数相等。
7. 设 n 是正整数。证明在任意一组 n 个连续的正整数中恰好有 1 个被 n 整除。
8. 证明如果 f 是从 S 到 T 的函数，其中 S 和 T 是有穷集，满足 $|S|>|T|$，那么在 S 中存在元素 s_1 和 s_2 使得 $f(s_1)=f(s_2)$，或者换句话说，f 不是一对一的。
9. 在一个大学里每个学生来自 50 个州中的一个州，那么必须有多少个学生注册才能保证至少有 100 个学生来自同一个州？

*10. 设 $(x_i, y_i)(i=1, 2, 3, 4, 5)$ 是 xy 平面上一组具有整数坐标的 5 个不同的点。证明至少有一对点的连线中点的坐标是整数。

*11. 设 $(x_i, y_i, z_i)(i=1, 2, 3, 4, 5, 6, 7, 8, 9)$ 是 xyz 空间中一组具有整数坐标的 9 个不同的点。证明至少有一对点的连线中点的坐标是整数。

12. 至少需要多少个有序对 (a, b) 才能保证存在两个有序对 (a_1, b_1) 和 (a_2, b_2) 使得 $a_1 \textbf{ mod } 5=a_2 \textbf{ mod } 5$ 并且 $b_1 \textbf{ mod } 5=b_2 \textbf{ mod } 5$。
13. a)如果从前 8 个正整数中选 5 个整数一定存在一对整数其和等于 9。
    b)如果不是选 5 个而是选 4 个整数，a)的结论还为真吗？
14. a)如果从前 10 个正整数中选 7 个整数，一定至少存在 2 对整数其和等于 11。
    b)如果不是选 7 个而是选 6 个整数，a)的结论还为真吗？
15. 从集合{1, 2, 3, 4, 5, 6}中必须选多少个数才能保证其中至少有一对数之和等于 7？
16. 从集合{1, 3, 5, 7, 9, 11, 13, 15}中必须选多少个数才能保证其中至少有一对数之和等于 16？
17. 一个公司在仓库储存产品。仓库中的存储柜由它们的通道、在通道中的位置和货架来指定。整个仓库有 50 个通道，每个通道 85 个水平位置，每个位置 5 个货架。公司产品数至少是多少才能使得在同一个存储柜中至少有 2 个产品？
18. 设一个小学院的离散数学班中有 9 个学生
    a)证明这个班一定至少有 5 个男生，或者至少有 5 个女生。
    b)证明这个班一定至少有 3 个男生，或者至少有 7 个女生。
19. 在 25 个学生的离散数学班中，学生有一年级的、二年级的或者三年级的。
    a)证明这个班至少有 9 个是一年级的，或至少有 9 个是二年级的，或至少有 9 个是三年级的。
    b)证明这个班至少有 3 个是一年级的，或至少有 19 个是二年级的，或至少有 5 个是三年级的。
20. 在序列 22，5，7，2，23，10，15，21，3，17 中找出一个最长的递增子序列和一个最长的递减子序列。
21. 构造 16 个正整数的序列，使得它没有 5 项的递增或递减子序列。
22. 如果 101 个不同高度的人站在一条线上，证明可能找到 11 个人使得他们站在线上的高度是按递增或者递减顺序。

*23. 用伪码描述一个算法产生一个不同整数序列的最大递增或递减子序列。

24. 证明在任一组 5 个人中(其中任两个人或者是朋友或者是敌人)不一定有 3 个人彼此都是朋友或者 3 个

人彼此都是敌人。

25. 证明在任一组 10 个人中(其中任两个人或者是朋友或者是敌人)或存在 3 个人彼此都是朋友，或存在 4 个人彼此都是敌人；并且存在 3 个人彼此是敌人，或存在 4 个人彼此是朋友。
26. 使用练习 25 证明在任一组 20 个人中(其中任两个人或者是朋友或者是敌人)或存在 4 个人彼此都是朋友，或存在 4 个人彼此都是敌人。
27. 证明如果 n 是正整数，$n\geqslant 2$，那么拉姆齐数 $R(2, n)$ 等于 n。
28. 证明如果 m 和 n 是正整数，$m\geqslant 2$，$n\geqslant 2$，那么拉姆齐数 $R(m, n)$ 和 $R(n, m)$ 相等。
29. 证明在加利福尼亚州(人口 3400 万)至少有 6 个人姓名的 3 个缩写字母相同并且他们生在一年的同一天(但不一定是同一年)。
30. 证明如果美国工薪阶层有 100 000 000 人的工资低于 1 000 000 美元，那么去年有 2 个人挣的钱恰好相同(准确到美分)。
31. 一个大学有 38 个不同的时间段来安排课程，如果有 677 门不同的课程，那么需要多少个不同的教室？
32. 一个计算机网络由 6 台计算机组成。每台计算机至少直接连接到一台其他的计算机。证明网络中至少有两台计算机直接连接相同数目的其他计算机。
33. 一个计算机网络由 6 台计算机组成。每台计算机直接连接到零台或者更多的其他计算机。证明网络中至少有两台计算机直接连接相同数目的其他计算机。
34. 把 8 台计算机连接到 4 台打印机上，为保证 4 台计算机可以直接访问 4 台不同的打印机，找出至少需要多少条缆线。证明你的答案。
35. 把 100 台计算机连接到 20 台打印机上，为保证 20 台计算机可以直接访问 20 台不同的打印机，找出至少需要多少条缆线。证明你的答案。

*36. 证明在至少 2 个人的聚会中，存在 2 个人认识人数相同的其他人。

37. 一个摔跤选手是 75 小时之内的冠军。该选手一小时至少赛一场，但总共不超过 125 场。证明存在着连续的若干个小时使得该选手恰好进行了 24 场比赛。

*38. 如果在练习 37 中的 24 替换如下，命题是否为真？

a)2　　b)23　　c)25　　d)30

39. 如果 f 是从 S 到 T 的函数，其中 S 和 T 是有穷集，并且 $m=\lceil |S|/|T|\rceil$，那么证明至少存在 S 的 m 个元素映射到 T 的同一个值。即存在 S 中的元素 $s_1, s_2, \cdots, s_m$ 使得 $f(s_1)=f(s_2)=\cdots=f(s_m)$。
40. 一条街道上有 51 所房子，每所房子的地址在 1000 到 1099 之间(包括 1000 与 1099)，证明至少有 2 所房子的地址是连续的。

*41. 设 x 是无理数。证明对于某个不超过 n 的正整数 j，在 jx 和到 jx 最近的整数之间的差的绝对值小于 $1/n$。

42. 设 $n_1, n_2, \cdots, n_t$ 是正整数。证明如果 $n_1+n_2+\cdots+n_t-t+1$ 个物体放到 t 个盒子里，则对某个 i，$i=1, 2, \cdots, t$，第 i 个盒子包含了至少 n_i 个物体。

*43. 在这个问题中概述了基于广义鸽巢原理的定理 3 的证明，使用的记号与教科书中的证明一样。

a)假定 $i_k\leqslant n$，$k=1, 2, \cdots, n^2+1$。使用广义鸽巢原理证明存在 $n+1$ 个项 $a_{k_1}, a_{k_2}, \cdots, a_{k_{n+1}}$ 满足 $i_{k_1}=i_{k_2}=\cdots=i_{k_{n+1}}$，其中 $1<k_1<k_2<\cdots<k_{n+1}$。

b)证明 $a_{k_j}>a_{k_{j+1}}$，$j=1, 2, \cdots, n$。[提示：假定 $a_{k_j}<a_{k_{j+1}}$，证明这将推出 $i_{k_j}>i_{k_{j+1}}$ 的矛盾。]

c)使用(a)和(b)证明如果没有长为 $n+1$ 的递增子序列，那么一定有同样长的递减子序列。

4.3 排列与组合

4.3.1 引言

假定一个网球队有 10 个选手，教练必须选择 5 个选手外出到另一个学校参加比赛。此外，教练还必须准备 4 个选手的排名表来进行 4 场单打比赛。在这一节，将提出一些方法来计数无序挑选 5 名外出比赛选手的各种组合和参加 4 场单打比赛的不同的 4 个选手的排名表。更一般地，

将引入计数一个有穷集的不同个体的无序选择和有序安排的技术。

4.3.2 排列

不同个体集合的一个排列是这些个体的一种有序安排。我们也对集合的某些元素的有序安排感兴趣。一个集合的 r 个元素的有序安排叫做 r 排列。

例1 设 $S=\{1, 2, 3\}$。3，1，2 是 S 的一个排列，3，2 是 S 的一个 2 排列。■

一个 n 元集的 r 排列数记为 $P(n, r)$。我们可以使用乘积法则求出 $P(n, r)$。

定理1 具有 n 个不同元素的集合的 r 排列数是

$$P(n,r) = n(n-1)(n-2)\cdots(n-r+1)$$

证 选择这个排列的第一个元素可以有 n 种方法，因为集合中有 n 个元素。选择排列的第二个元素有 $n-1$ 种方法，由于在使用了为第一个位置挑出的元素之后集合里还留下了 $n-1$ 个元素。类似地，选择第三个元素有 $n-2$ 种方法，依此类推，直到选择第 r 个元素恰好有 $n-(r-1)=n-r+1$ 种方法。因此，由乘积法则，存在

$$n(n-1)(n-2)\cdots(n-r+1)$$

个集合的 r 排列。□

从定理1得出

$$P(n,r) = n(n-1)(n-2)\cdots(n-r+1) = n!/(n-r)!$$

特别地有 $P(n, n)=n!$。我们将用一些例子说明这个结果。

例2 在进入竞赛的 100 个不同的人中有多少种方法选出一个一等奖得主、一个二等奖得主和一个三等奖得主？

解 不管哪个人得哪个奖，选取 3 个得奖人的方法数是从 100 个元素的集合中有序选择 3 个元素的方法数，即 100 个元素的集合的 3-排列数。因此，答案是

$$P(100,3) = 100 \cdot 99 \cdot 98 = 970\,200$$ ■

例3 假定有 8 个赛跑运动员。第一名得到一枚金牌，第二名得到一枚银牌，第三名得到一枚铜牌。如果比赛可能出现所有可能的结果，有多少种不同的颁奖方式？

解 颁奖方式就是 8 元素的集合的 3-排列数。因此存在 $P(8, 3)=8\cdot 7\cdot 6=336$ 种可能的颁奖方式。■

例4 假定一个女推销员要访问 8 个不同的城市。她的访问必须从某个指定的城市开始，但对其他 7 个城市的访问可以按照任何次序进行。当访问这些城市时，这个女推销员可以有多少种可能的次序？

解 由于第一个城市是确定的，而其他 7 个城市可以是任意的顺序，故城市之间可能的路径数是 7 个元素的排列数。因此，这个女推销员有 $7!=7\cdot 6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1=5040$ 种方式选择她的旅行。比如说，如果这个女推销员想要在城市中找出具有最短距离的路径，并且她对每一条可能的路径计算总距离，那么她必须考虑 5040 条路径。■

例5 字母 $ABCDEFGH$ 有多少种排列包含串 ABC？

解 由于字母 ABC 必须成组出现，我们可以通过找 6 个对象，即组 ABC 和单个字母 D、E、F、G 和 H 的排列数得到答案。由于这 6 个对象可以按任何次序出现，存在 $6!=720$ 种 $ABCDEFGH$ 字母的排列，其中 ABC 成组出现。■

4.3.3 组合

集合元素的一个 r 组合是从这个集合无序选取的 r 个元素。于是，简单地说，一个 r 组合是这个集合的一个 r 个元素的子集。

例 6　设 S 是集合 $\{1, 2, 3, 4\}$，那么 $\{1, 3, 4\}$ 是 S 的一个 3 组合。■

具有 n 个不同元素集合的 r 组合数记为 $C(n, r)$。注意 $C(n, r)$ 也记作 $\binom{n}{r}$ 并且称为二项式系数。在 4.4 节我们将学习这个记号。

例 7　因为 $\{a, b, c, d\}$ 的 2-组合是 $\{a, b\}$，$\{a, c\}$，$\{a, d\}$，$\{b, c\}$，$\{b, d\}$ 和 $\{c, d\}$ 6 个子集，我们有 $C(4, 2)=6$。■

我们可以用关于集合的 r 排列数的公式确定 n 元素的集合的 r 组合数。为此只需注意到集合的 r 排列可以按下述方法得到：首先构成集合的 r 组合，接着排列这些组合中的元素。下面的定理给出了 $C(n, r)$ 的值，它的证明就是基于这个观察。

定理 2　设 n 是正整数，r 是满足 $0 \leqslant r \leqslant n$ 的整数，n 元素的集合的 r 组合数等于

$$C(n,r) = \frac{n!}{r!(n-r)!}$$

证　可以如下得到这个集合的 r 排列。先构成集合的 $C(n, r)$ 个 r 组合，然后以 $P(n, r)$ 种方式排序每个 r 组合中的元素，这可以用 $P(r, r)$ 种方式来做。因此，

$$P(n,r) = C(n,r) \cdot P(r,r)$$

这就推出

$$C(n,r) = \frac{P(n,r)}{P(r,r)} = \frac{n!/(n-r)!}{r!/(r-r)!} = \frac{n!}{r!(n-r)!}$$

□

注意　定理 2 中的公式，尽管很清楚，但对大的 n 和 r 的值计算 $C(n, r)$ 时也没有用。因为实际上只能对小的整数值来计算阶乘的准确值，并且当使用浮点数算数时，定理 2 的公式可能产生一个不是整数的值。在实践中，为计算 $C(n, r)$ 你可以从分子和分母中消去分母的大因式中所有的项，把分子中所有没有消去的项相乘，然后除以分母中较小的因式。许多计算器有一个关于 $C(n, r)$ 的内置函数。

推论 1 给出了一个关于集合 r 组合数的有用的恒等式。

推论 1　设 n 和 r 是满足 $r \leqslant n$ 的非负整数，那么 $C(n, r)=C(n, n-r)$。

证　由定理 2 得到

$$C(n,r) = \frac{n!}{r!(n-r)!}$$

和

$$C(n,n-r) = \frac{n!}{(n-r)![n-(n-r)]!} = \frac{n!}{(n-r)!r!}$$

因此，$C(n, r)=C(n, n-r)$。□

通过证明在推论 1 的等式两边使用了不同的理由对同样的个体进行计数，也可以证明推论 1。我们在定义 1 描述了这种重要的证明类型。

定义 1　组合证明是一种证明，在这个证明中使用计数的论述而不使用某些其他的方法如代数技巧来证明一个定理。

可以使用组合证明证出许多涉及二项式系数的恒等式。如果可以显示一个恒等式两边通过不同的方法计数了同样的元素，那么对这个恒等式就可以使用组合证明。现在提供一个推论 1 的组合证明。

证　设 S 是 n 个元素的集合。S 的每个具有 r 个元素的子集 A 对应于一个具有 $n-r$ 个元素的子集，即 $\overline{A}$，因此 $C(n, r)=C(n, n-r)$。□

例 8　有多少种方式从 10 个选手的网球队中选择 5 个选手外出参加在另一个学校的比赛？

解 答案由 10 元素集合的 5-组合数给出。根据定理 2，这个组合数是

$$C(10,5)=\frac{10!}{5!5!}=252。$$ ■

例 9 一组 30 个人被培训作为宇航员去完成首次登陆火星的任务。有多少种方式选出 6 个人的小组来完成这个任务（假设所有的小组成员有同样的工作）？

解 因为不考虑这些人被选的次序，从 30 个人中选 6 个人的小组的方式数是 30 元素集合的 6-组合数。根据定理 2，这种组合数是

$$C(30,6)=\frac{30!}{6!24!}=\frac{30\cdot 29\cdot 28\cdot 27\cdot 26\cdot 25}{6\cdot 5\cdot 4\cdot 3\cdot 2\cdot 1}=593\ 775$$ ■

例 10 有多少个长为 n 的二进制串恰好包含 r 个 1？

解 在长为 n 的二进制串中 r 个 1 的位置构成了集合{1, 2, …, n}的 r 组合。因此，有 $C(n, r)$个长为 n 的二进制串恰好包含 r 个 1。 ■

例 11 为开发学校的离散数学课程要选出一个委员会。如果数学系有 9 个教师，计算机科学系有 11 个教师。而这个委员会要由 3 个数学系的教师和 4 个计算机科学系的教师组成，那么有多少种选择方式？

解 由乘积法则，答案是 9 元素集合的 3 组合数与 11 元素集合的 4 组合数之积。根据定理 2，选择这个委员会的方式数是

$$C(9,3)\cdot C(11,4)=\frac{9!}{3!6!}\cdot\frac{11!}{4!7!}=84\cdot 330=27\ 720$$ ■

练习

1. 列出{a, b, c}的所有排列。
2. 集合{a, b, c, d, e, f, g}有多少个排列？
3. {a, b, c, d, e, f, g}有多少个排列以 a 结尾？
4. 令 S={1, 2, 3, 4, 5}，
   a)列出 S 的所有 3 排列。
   b)列出 S 的所有 3 组合。
5. 求出下面的每个值
   a)$P(6, 3)$ b)$P(6, 5)$ c)$P(8, 1)$
   d)$P(8, 5)$ e)$P(8, 8)$ f)$P(10, 9)$
6. 求出下面的每个值
   a)$C(5, 1)$ b)$C(5, 3)$ c)$C(8, 4)$
   d)$C(8, 8)$ e)$C(8, 0)$ f)$C(12, 6)$
7. 求出 9 元素集合的 5 排列数。
8. 如果不允许并列名次，在结束比赛时 5 个赛跑运动员有多少种不同的排名次序？
9. 在一场 12 匹马的赛马中，如果所有的比赛结果都是可能的，对于第一名、第二名和第三名有多少种可能性？
10. 有 6 个不同的人竞选州长。有多少种不同的次序在选票上打印竞选者的名字？
11. 多少个 10 位二进制串包含
    a)恰好 4 个 1？ b)至多 4 个 1？ c)至少 4 个 1？ d)0 的个数和 1 的个数相等？
12. 多少个 12 位二进制串包含
    a)恰好 3 个 1？ b)至多 3 个 1？ c)至少 3 个 1？ d)0 的个数和 1 的个数相等？
13. 一个组有 n 个男士和 n 个女士。如果把他们男女相间地排成一排，有多少种方式？
14. 有多少种不同的方式选择两个小于 100 的正整数？
15. 有多少种不同的方式从英语字母表中选择 5 个字母？

16. 一个 10 元素集合有多少个子集含有奇数个元素？
17. 一个 100 个元素的集合有多少个子集包含的元素多于 2 个？
18. 一个硬币被掷 8 次，每次可能出现头像或者非头像。有多少种可能的结果
   a)包含各种不同的情况？　b)包含恰好 3 个头像？
   c)包含至少 3 个头像？　d)头像和非头像的数目相等？
19. 一个硬币被掷 10 次，每次可能出现头像或者非头像。有多少种可能的结果
   a)包含各种不同的情况？　b)包含恰好 2 个头像？
   c)至多有 3 个不是头像？　d)头像和非头像的数目相等？
20. 多少个 10 位二进制串
   a)恰好有 3 个 0？　b)0 比 1 多？　c)至少有 7 个 1？　d)至少有 3 个 1？
21. 字母 $ABCDEFG$ 有多少个排列包含
   a)串 BCD？　b)串 $CFGA$？　c)串 BA 和 GF？　d)串 ABC 和 DE？
   e)串 ABC 和 CDE？　f)串 CBA 和 BED？
22. 字母 $ABCDEFGH$ 有多少个排列包含
   a)串 ED？　b)串 CDE？　c)串 BA 和 FGH？　d)串 AB、DE 和 GH？
   e)串 CAB 和 BED？　f)串 BCA 和 ABF？
23. 有多少种方式使得 8 个男士和 5 个女士站成一排并且没有两个女士彼此相邻？（提示：先放男士，然后考虑女士可能的位置。）
24. 有多少种方式使得 10 个女士和 6 个男士站成一排并且没有两个男士彼此相邻？（提示：先放女士，然后考虑男士可能的位置。）
25. 把编号为 1，2，…，100 的 100 张票卖给 100 个不同的人来抽奖。有 4 项不同的奖，包括 1 项大奖（到塔希提岛旅游）。如果满足下面的条件，有多少种不同的抽奖方式？
   a)没有限制。
   b)拿 47 号票的人赢了大奖。
   c)拿 47 号票的人赢了一项奖。
   d)拿 47 号票的人没赢奖。
   e)拿 19 和 47 号票的人都赢了奖。
   f)拿 19、47 和 73 号票的人都赢了奖。
   g)拿 19、47、73 和 97 号票的人都赢了奖。
   h)拿 19、47、73 和 97 号票的人都没赢奖。
   i)拿 19、47、73 或 97 号票的人赢了大奖。
   j)拿 19 和 47 号票的人赢了奖，但拿 73 和 97 号票的人没赢奖。
26. 一个垒球队的 13 个人出席一场比赛。
   a)有多少种方式选 10 个选手上场？
   b)有多少种方式从 13 个在场的人中分配 10 个选手的位置？
   c)13 个出席的人中有 3 个女士。如果上场的选手中要求至少有一个女士，那么有多少种方式选择 10 个选手？
27. 一个俱乐部有 25 个成员。
   a)有多少种方式从中选择 4 个人作为董事会成员。
   b)有多少种方式从中选出俱乐部的主席、副主席、书记和司库？
28. 一个教授写了 40 道离散数学的真假判定题。在这些题中有 17 个语句为真。如果可以按照任意次序排列这些题，可能有多少种不同的答案？
29. 用不超过 100 的正整数构成 4 排列，其中有多少个排列包含 3 个连续的整数？
   a)这里的连续指按照整数通常的顺序，并且这些连续整数可能被排列中的其他整数分开。
   b)这里的连续不但指整数是连续的，而且它们在排列中的位置也是连续的。

30. 一所学校的数学系有7名女教师和9名男教师。

a)有多少种方式从中选出5人的委员会并使其中包含至少1名女教师?

b)有多少种方式从中选出5人的委员会并使其中包含至少1名女教师和至少1名男教师。

31. 英语字母表中包含21个辅音和5个元音。由英语字母表的6个小写字母可构成多少字符串使得它们包含

a)恰好1个元音? b)恰好2个元音? c)至少1个元音? d)至少2个元音?

32. 由英语字母表中的6个小写字母可构成多少字符串使得它们包含

a)字母 a?

b)字母 a 和 b?

c)字母 a 和 b,其中 a 在 b 前边的邻接位置,同时所有的字母都不相同?

d)字母 a 和 b,其中 a 在 b 左边的某个位置,同时所有的字母都不相同?

33. 假定某个系包含10名男士和15名女士。有多少种方式组成一个6人委员会且使得它含有相同数量的男士和女士?

34. 假定某个系包含10名男士和15名女士。有多少种方式组成一个6人委员会且使得它含有的女士比男士多?

35. 有多少个二进制串恰好包含8个0和10个1,如果每个0后面紧跟着1个1?

36. 有多少个二进制串恰好包含5个0和14个1,如果每个0后面紧跟着2个1?

37. 有多少个10位二进制串包含至少3个1和至少3个0?

38. 有多少种方式从联合国中选择12个国家成为理事国且使得3个选自45个国家的一组,4个选自57个国家的一组,其他的选自剩下的69个国家?

39. 有多少种方式用3个字母后跟3个数字组成汽车牌照并且没有字母和数字出现2次?

40. 6个人围着一个圆桌就座,如果通过旋转圆桌而使得就座方式从一种变成另一种,就认为它们是同一种方式,那么有多少种不同的就座方式?

41. 如果允许出现并列名次,3匹马参加马赛有多少种结果?(注:可以2匹或3匹马并列。)

*42. 如果允许并列名次,4匹马参加马赛有多少种结果?(注:由于允许并列名次,4匹马中多少匹并列都是可能的。)

*43. 有6名运动员参加百米赛跑。如果允许并列名次,有多少种方式授予3块奖牌?(跑得最快的运动员得金牌,恰好只被一个运动员超过的运动员得银牌,恰好被2个运动员超过的运动员得铜牌。)

*44. 为避免世界杯足球锦标赛总决赛中出现并列名次,通常采用下述过程:每个队按照预定的顺序选出5名球员。每名球员罚一个球,第一队的球员先罚,接着第二队的球员再罚,依照指定的顺序依次交替罚球。如果在10次罚球后得分还相等,再次重复这个过程。如果在20次罚球后得分仍旧相等,进行加赛时间的射门,第一个得分的队得胜。

a)如果比赛进行第一轮的10个罚球,并且这轮比赛结束时一个队不可能与另一个队得分相等,那么有多少种不同的得分场面?

b)如果比赛进行第二轮的10个罚球,对第一和第二轮罚球可能有多少种不同的得分场面?

c)如果比赛在两轮每队罚5个球的加赛以后最多再射门10次,那么整个加赛过程可能有多少种得分场面?

4.4 二项式系数

正如我们在4.3节谈到的,具有 n 个元素的集合的 r-组合数常常记作 $\binom{n}{r}$,由于这些数出现在二项式的幂 $(a+b)^n$ 的展开式中作为系数,因此这些数叫做**二项式系数**。我们将讨论二项式定理,这个定理将二项式的幂表示成与二项式系数有关的项之和。我们将用组合证明来证明这个定理。我们也将显示怎样用组合证明来建立某些恒等式,它们是表示二项式系数之间关系的许

多不同恒等式中的一部分。

4.4.1 二项式定理

二项式定理给出了二项式幂的展开式的系数。一个二项式只不过是两项的和，例如 $x+y$。（这些项可以是常数与变量的积，但这里先不考虑。）例 1 说明了为什么这个定理成立。

例 1 $(x+y)^3$ 的展开式可以使用组合推理而不是用三个项的乘法得到。当 $(x+y)^3=(x+y)(x+y)(x+y)$ 被展开时，把所有由第一个和的一项、第二个和的一项与第三个和的一项产生的乘积加起来。从而出现了形如 x^3、x^2y、xy^2 和 y^3 的项。为得到形如 x^3 的项，在每个和里必须选择一个 x，只有一种方式能做到这一点。因此，乘积中 x^3 项的系数是 1。为得到形如 x^2y 的项，必须从三个和中的两个和里选 x（而因此在另一个和里选 y）。于是，这种项的个数是三个个体的 2-组合数，即 $\binom{3}{2}$。类似地，形如 xy^2 项的个数是三个和中选一个来提供 x 的方式数（而另两个和中都要选 y），有 $\binom{3}{1}$ 种方式能够做到这一点。最后，得到 y^3 的唯一方式是三个和的每一个都选择 y，恰好有一种方式能够做到这一点。因此得到

$$
\begin{aligned}
(x+y)^3 &= (x+y)(x+y)(x+y) = (xx+xy+yx+yy)(x+y) \\
&= xxx+xxy+xyx+xyy+yxx+yxy+yyx+yyy \\
&= x^3+3x^2y+3xy^2+y^3
\end{aligned}
$$

■

现在叙述二项式定理。

定理 1 *二项式定理* 设 x 和 y 是变量，n 是非负整数，那么

$$
\begin{aligned}
(x+y)^n &= \sum_{j=0}^{n}\binom{n}{j}x^{n-j}y^j \\
&= \binom{n}{0}x^n+\binom{n}{1}x^{n-1}y+\binom{n}{2}x^{n-2}y^2+\cdots+\binom{n}{n-1}xy^{n-1}+\binom{n}{n}y^n
\end{aligned}
$$

证 这里给出定理的组合证明。当乘积被展开时其中的项都是下述形式：$x^{n-j}y^j(j=0,1,2,\cdots,n)$。为计数形如 $x^{n-j}y^j$ 的项数，观察到必须从 n 个和中选 $n-j$ 个 x（从而乘积中其他的 j 个项都是 y）才能得到这种项。因此，$x^{n-j}y^j$ 的系数是 $\binom{n}{n-j}$，它等于 $\binom{n}{j}$，定理得证。 □

例 2～例 4 说明了二项式定理的应用。

例 2 $(x+y)^4$ 的展开式是什么？

解 由二项式定理得到

$$
\begin{aligned}
(x+y)^4 &= \sum_{j=0}^{4}\binom{4}{j}x^{4-j}y^j \\
&= \binom{4}{0}x^4+\binom{4}{1}x^3y+\binom{4}{2}x^2y^2+\binom{4}{3}xy^3+\binom{4}{4}y^4 \\
&= x^4+4x^3y+6x^2y^2+4xy^3+y^4
\end{aligned}
$$

■

例 3 在 $(x+y)^{25}$ 的展开式中 $x^{12}y^{13}$ 的系数是什么？

解 由二项式定理得到这个系数是

$$\binom{25}{13}=\frac{25!}{13!12!}=5\ 200\ 300$$

■

例 4 在$(2x-3y)^{25}$的展开式中$x^{12}y^{13}$的系数是什么？

解 首先注意到这个表达式等于$(2x+(-3y))^{25}$。由二项式定理，我们有

$$(2x+(-3y))^{25}=\sum_{j=0}^{25}\binom{25}{j}(2x)^{25-j}(-3y)^j$$

因此，当$j=13$时得到展开式中$x^{12}y^{13}$的系数，即

$$\binom{25}{13}2^{12}(-3)^{13}=-\frac{25!}{13!12!}2^{12}3^{13}$$ ■

我们可以用二项式定理证明某些有用的恒等式。正如推论 1、2 和 3 所示。

推论 1 设n为非负整数，那么

$$\sum_{k=0}^{n}\binom{n}{k}=2^n$$

证 用二项式定理，令$x=1$和$y=1$，我们有

$$2^n=(1+1)^n=\sum_{k=0}^{n}\binom{n}{k}1^k1^{n-k}=\sum_{k=0}^{n}\binom{n}{k}$$

这正是所需要的结果。□

推论 1 也有一个好的组合证明，我们现在给出这个证明。

证 一个n元素集合有2^n个不同的子集。每个子集有 0 个元素、1 个元素、2 个元素，…，或n个元素。具有 0 个元素的子集有$\binom{n}{0}$个，1 个元素的子集有$\binom{n}{1}$个，2 个元素的子集有$\binom{n}{2}$个，…，n个元素的子集有$\binom{n}{n}$个。于是

$$\sum_{k=0}^{n}\binom{n}{k}$$

计数了n元素集合的子集总数。这证明了

$$\sum_{k=0}^{n}\binom{n}{k}=2^n$$ □

推论 2 设n是正整数，那么

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k}=0$$

证 由二项式定理得出

$$0=0^n=((-1)+1)^n=\sum_{k=0}^{n}\binom{n}{k}(-1)^k1^{n-k}=\sum_{k=0}^{n}\binom{n}{k}(-1)^k.$$

从而证明了推论。□

注意 推论 2 推出

$$\binom{n}{0}+\binom{n}{2}+\binom{n}{4}+\cdots=\binom{n}{1}+\binom{n}{3}+\binom{n}{5}+\cdots$$

推论 3 设n是非负整数，那么

$$\sum_{k=0}^{n}2^k\binom{n}{k}=3^n$$

证 注意到这个公式的左边是二项式定理提供的对$(1+2)^n$的展开式。因此，由二项式定理可以看出

$$(1+2)^n = \sum_{k=0}^{n}\binom{n}{k}1^{n-k}2^k = \sum_{k=0}^{n}\binom{n}{k}2^k$$

因此

$$\sum_{k=0}^{n} 2^k\binom{n}{k} = 3^n$$

□

4.4.2　帕斯卡恒等式和三角形

二项式系数满足许多不同的恒等式。现在我们介绍其中最重要的一些恒等式。

定理 2　帕斯卡恒等式[⊖]设 n 和 k 是满足 $n \geqslant k$ 的正整数，那么有

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$$

证　假定 T 是包含 $n+1$ 个元素的集合。令 a 是 T 的一个元素且 $S=T-\{a\}$。注意到 T 的包含 k 个元素的子集有 $\binom{n+1}{k}$ 个。然而 T 的含 k 个元素的子集或者包含 a 和 S 中的 $k-1$ 个元素，或者不包含 a 但包含 S 中的 k 个元素。由于 S 的 $k-1$ 元子集有 $\binom{n}{k-1}$ 个，故 T 含 a 在内的 k 元子集有 $\binom{n}{k-1}$ 个。又由于 S 的 k 元子集有 $\binom{n}{k}$ 个，故 T 的不含 a 的 k 元子集有 $\binom{n}{k}$ 个。从而得到

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$$

□

注意　这里给出了帕斯卡恒等式的一个组合证明。也可以从关于 $\binom{n}{r}$ 的公式通过代数推导来证明这个恒等式(见节末的练习 19)。

注意　对所有整数 n，可以用帕斯卡恒等式和初始条件 $\binom{n}{0}=\binom{n}{n}=1$ 递归地定义二项式系数。这些递归定义用于计算二项式系数，因为使用这些递归定义只需要整数加法。

帕斯卡恒等式是二项式系数以三角形表示的几何排列的基础，如图 4-6 所示。

这个三角形的第 n 行由二项式系数

$$\binom{n}{k}, k = 0,1,\cdots,n$$



组成。这个三角形叫做帕斯卡三角形。帕斯卡恒等式证明，当这个三角形中两个相邻的二项式系数相加时，就产生了下一行在这两个系数之间的二项式系数。



⊖ **布莱斯·帕斯卡(Blaise Pascal, 1623—1662)**　帕斯卡在幼年时就显现出他的才能，虽然他的父亲(对解析几何有过多项建树)为了鼓励他在其他方面的兴趣，不让他接触数学书。帕斯卡 16 岁时发现了圆锥曲线中的一个重要结果。18 岁时他设计了一部计算机，建造后将其卖出。帕斯卡和费马一道奠定了现代概率论的基础。在他的工作中有对现今称为帕斯卡三角形的一些发现。1654 年帕斯卡放弃了对数学的追求，转而研究神学。在那以后，他只有一次重返数学。一个晚上，他因剧烈牙痛而心烦意乱。他想通过研究摆线性质来缓解疼痛。不可思议的是牙疼居然减退了，他把这一点看成上天赞成他研究数学的暗示。

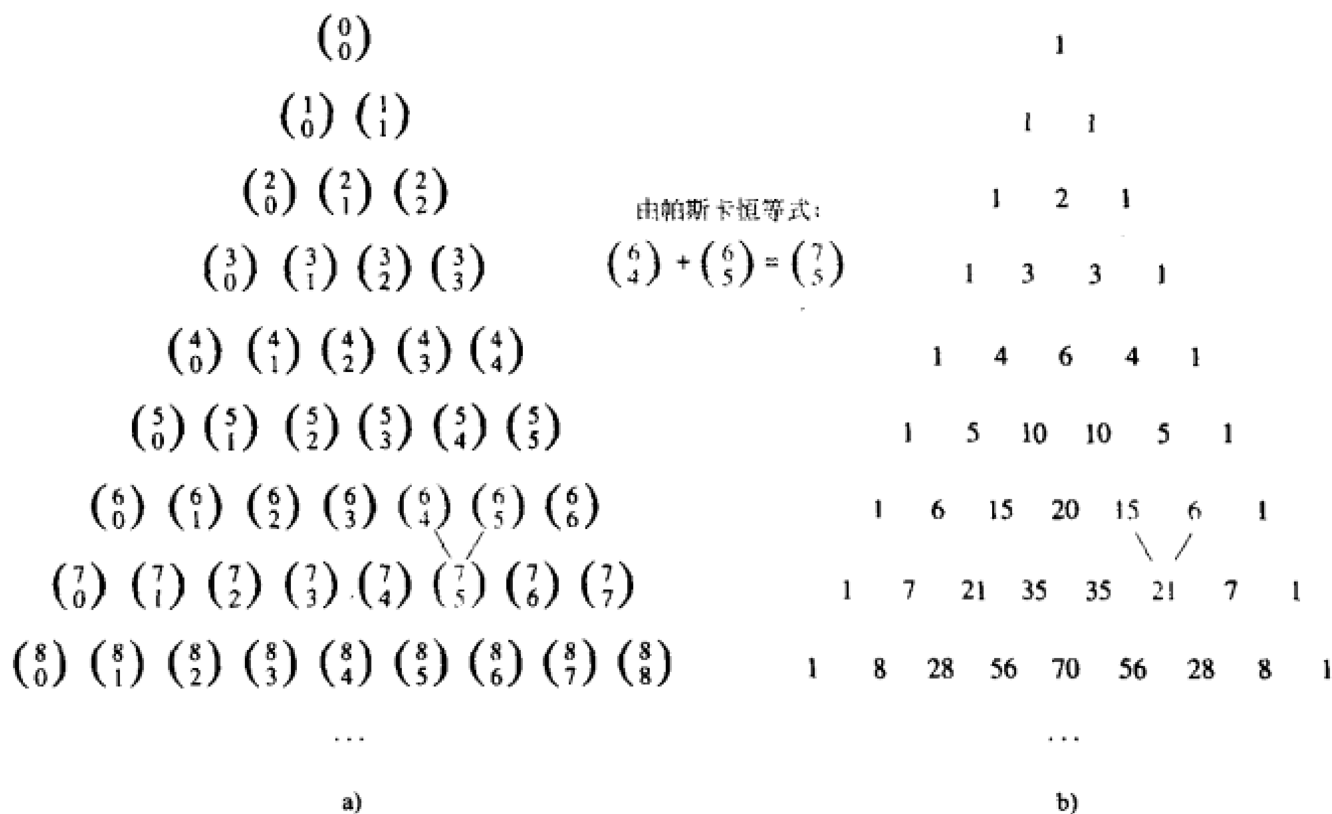


图 4-6 帕斯卡三角形

4.4.3 其他的二项式系数恒等式

我们从众多二项式系数恒等式中选两个恒等式，用它们的组合证明来作为本节的结束。

定理 3 *范德蒙德恒等式* 设 m、n 和 r 是非负整数，其中 r 不超过 m 或 n，那么

$$\binom{m+n}{r}=\sum_{k=0}^{r}\binom{m}{r-k}\binom{n}{k}$$

注意 这个恒等式是由 18 世纪的数学家亚历山大-舍费尔·范德蒙德[⊖]发现的。

证 假定在第一集合中有 m 个项，第二集合中有 n 个项。从这两个集合的并集中取 r 个元素的方式数是 $\binom{m+n}{r}$。从并集中取 r 个元素的另一种方式是先从第一个集合中取 k 个元素，接着从第二个集合中取 $r-k$ 个元素，其中 k 是满足 $0\leqslant k\leqslant r$ 的整数。由乘积法则，这可以用 $\binom{m}{k}\binom{n}{r-k}$ 种方式完成。所以，从这个并集中选取 r 个元素的总方式数等于

$$\binom{m+n}{r}=\sum_{k=0}^{r}\binom{m}{r-k}\binom{n}{k}$$

这就证明了范德蒙德恒等式。 □

推论 4 来自范德蒙德恒等式。

推论 4 如果 n 是一个非负整数，那么

⊖ **亚历山大-舍费尔·范德蒙德(Alexandre-Théophile Vandermonde，1735—1796)** 范德蒙德年幼时体弱多病，作医生的父亲让他从事音乐职业。但是后来，他对数学越来越感兴趣。他的完整的数学工作包含在 1771～1772 年发表的 4 篇论文中。这些论文包含了在方程求根、行列式理论以及骑士旅行问题(在 7.5 节的练习中介绍)的基础贡献。范德蒙德对数学的兴趣只持续了两年。后来，他在和声学、寒冷实验以及钢的制造等方面发表论文。他也对政治发生了兴趣，参加法国革命事业，并且在政府中担任了几个不同的职务。

$$\binom{2n}{n}=\sum_{k=0}^{n}\binom{n}{k}^2$$

证　在范德蒙德恒等式中令 $m=r=n$ 得到

$$\binom{2n}{n}=\sum_{k=0}^{n}\binom{n}{n-k}\binom{n}{k}=\sum_{k=0}^{n}\binom{n}{k}^2$$

最后一步的相等是使用了恒等式 $\binom{n}{k}=\binom{n}{n-k}$。□

我们可以通过计数具有不同性质的二进制串来证明组合恒等式，如定理 4 的证明所示。

定理 4　设 n 和 r 是非负整数，$r\leqslant n$，那么

$$\binom{n+1}{r+1}=\sum_{j=r}^{n}\binom{j}{r}$$

证　我们使用组合证明。由 4.3 节例 10，左边 $\binom{n+1}{r+1}$ 计数了 $n+1$ 位长包含了 $r+1$ 个 1 的二进制串。

我们证明，在具有 $r+1$ 个 1 的二进制串中，通过考虑与最后一个 1 可能位置的相关情况，等式右边计数了同样的对象。这最后一个 1 一定出现在位置 $r+1$，$r+2$，…，或者 $n+1$。此外，如果最后一个 1 出现在第 k 位，那么一定有 r 个 1 出现在前 $k-1$ 位。因此，根据 4.3 节例 10，这样的二进制串有 $\binom{k-1}{r}$ 个，对所有的 k 求和，其中 $r+1\leqslant k\leqslant n+1$，我们发现有

$$\sum_{k=r+1}^{n+1}\binom{k-1}{r}=\sum_{j=r}^{n}\binom{j}{r}$$

个 n 位二进制串恰含有 $r+1$ 个 1。（注意最后一步是改变变量 $j=k-1$ 的结果。）由于左边和右边计数了同样的对象，因此相等。这就完成了证明。□

练习

1. 求 $(x+y)^4$ 的展开式
   a)使用组合理由，如例 1 所示。　b)使用二项式定理。
2. 求 $(x+y)^5$ 的展开式
   a)使用组合理由，如例 1 所示。　b)使用二项式定理。
3. 求 $(x+y)^6$ 的展开式。
4. 求在 $(x+y)^{13}$ 的展开式中 x^5y^8 的系数。
5. 在 $(x+y)^{100}$ 的展开式中有多少个项？
6. 在 $(1+x)^{11}$ 中 x^7 的系数是什么？
7. 在 $(2-x)^{19}$ 中 x^9 的系数是什么？
8. 在 $(3x+2y)^{17}$ 中 x^8y^9 的系数是什么？
9. 在 $(2x-3y)^{200}$ 中 $x^{101}y^{99}$ 的系数是什么？

*10. 给出一个关于 $(x+1/x)^{100}$ 的展开式中 x^k 系数的公式，其中 k 是一个整数。

*11. 给出一个关于 $(x^2-1/x)^{100}$ 的展开式中 x^k 系数的公式，其中 k 是一个整数。

12. 帕斯卡三角形中包含二项式系数 $\binom{10}{k}$ $(0\leqslant k\leqslant 10)$ 的行是

1　10　45　120　210　252　210　120　45　10　1

用帕斯卡恒等式计算出在帕斯卡三角形中紧接这行下面的另一行。

13. 帕斯卡三角形中包含二项式系数$\binom{9}{k}(0\leqslant k\leqslant 9)$的行是什么？

14. 证明如果 n 是正整数，则

$$1=\binom{n}{0}<\binom{n}{1}<\cdots<\binom{n}{\lfloor n/2\rfloor}=\binom{n}{\lceil n/2\rceil}>\cdots>\binom{n}{n-1}>\binom{n}{n}=1$$

15. 证明对一切正整数 n 和 $k(0\leqslant k\leqslant n)$，$\binom{n}{k}\leqslant 2^n$。

16. a)用练习 14 和推论 1 证明如果 n 是大于 1 的整数，那么 $\binom{n}{\lfloor n/2\rfloor}\geqslant 2^n/n$。

b)从 a)确定如果 n 是正整数，那么$\binom{2n}{n}\geqslant 4^n/2n$。

17. 证明如果 n 和 k 是整数，其中 $1\leqslant k\leqslant n$，那么$\binom{n}{k}\leqslant n^k/2^{k-1}$。

18. 设 b 是整数，$b\geqslant 7$。使用二项式定理和帕斯卡三角形中适当的行找出$(11)_b^4$ 的以 b 为基的展开式[就是以 b 为基的数$(11)_b$ 的 4 次方]。

19. 使用关于$\binom{n}{r}$的公式证明帕斯卡恒等式。

20. 设 k 和 n 是整数，$1\leqslant k<n$，证明*六边形恒等式*，这些项在帕斯卡三角形中构成六边形。

$$\binom{n-1}{k-1}\binom{n}{k+1}\binom{n+1}{k}=\binom{n-1}{k}\binom{n}{k-1}\binom{n+1}{k+1}$$

21. 证明如果 n 和 k 是整数，$1\leqslant k\leqslant n$，那么

$$k\binom{n}{k}=n\binom{n-1}{k-1}$$

a)使用组合证明[*提示*：证明恒等式两边计数了从一个 n 元素集合中选 k 个元素，然后从这个子集中再选 1 个元素的方法]。

b)使用基于 4.3 节定理 2 给出的$\binom{n}{r}$公式的代数证明。

22. 证明恒等式$\binom{n}{r}\binom{r}{k}=\binom{n}{k}\binom{n-k}{r-k}$，其中 n、r 和 k 是非负整数且 $r\leqslant n$，$k\leqslant r$。

a)使用组合证明。

b)使用以 n 元素集合的 r-组合数公式为基础的论证。

23. 证明如果 n 和 k 是正整数，那么

$$\binom{n+1}{k}=(n+1)\binom{n}{k-1}/k$$

使用这个恒等式构造一个二项式系数的归纳的定义。

24. 证明如果 p 是素数，k 是满足 $1\leqslant k\leqslant p-1$ 的整数，那么 p 整除$\binom{p}{k}$。

25. 设 n 是正整数，证明$\binom{2n}{n+1}+\binom{2n}{n}=\binom{2n+2}{n+1}/2$。

*26. 设 n 和 k 是整数，$1\leqslant k\leqslant n$，证明$\sum_{k=1}^{n}\binom{n}{k}\binom{n}{k-1}=\binom{2n+2}{n+1}/2-\binom{2n}{n}$。

*27. 证明

$$\sum_{k=0}^{r}\binom{n+k}{k}=\binom{n+r+1}{r}$$

其中 n 和 r 是正整数。

a)用组合论证。　　b)用帕斯卡恒等式。

28. 证明如果 n 是正整数则 $\binom{2n}{2}=2\binom{n}{2}+n^2$。

a)使用组合论证。　　b)通过代数推导。

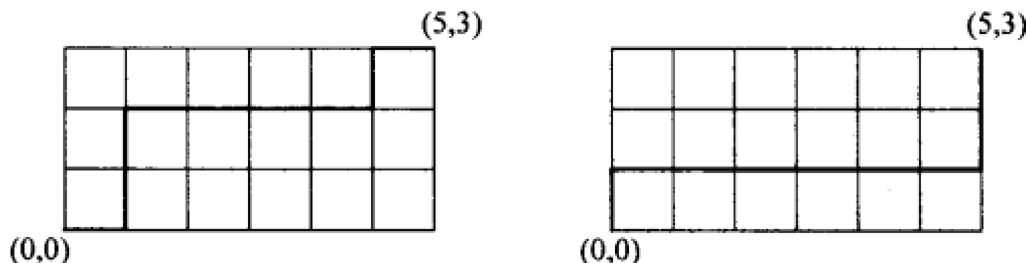
*29. 给出关于 $\sum_{k=1}^{n}k\binom{n}{k}=n2^{n-1}$ 的组合证明。[提示：以两种方法计数选择一个委员会然后选择这个委员会领导的方式数。]

*30. 给出关于 $\sum_{k=1}^{n}k\binom{n}{k}^2=n\binom{2n-1}{n-1}$ 的组合证明。[提示：用两种方法计数选择一个委员会的方式数，如果这个委员会有 n 个成员，要求这些成员选自 n 个数学教授和 n 个计算机科学教授，并使得委员会的主席是数学教授。]

31. 证明一个非空集合具有奇数个元素的子集数与具有偶数个元素的子集数相等。

*32. 使用数学归纳法证明二项式定理。

33. 在这个练习里我们将要计数 xy 平面上在原点和 (m, n) 点之间的路径数。这些路径由一系列步构成，其中每一步是向右或者向上移动一个单位(不允许向左或向下移动)。下图给出了两条这种从(0, 0)到(5, 3)的路径(用粗线标识)。



a)证明上述每条这种类型的路径可以用由 m 个 0 和 n 个 1 组成的二进制串表示，其中 0 表示向右移动一个单位，1 表示向上移动一个单位。

b)从 a)推断存在着 $\binom{m+n}{n}$ 条所求类型的路径。

34. 用练习 33 证明 $\binom{n}{k}=\binom{n}{n-k}$，其中 k 是一个整数，满足 $0\leqslant k\leqslant n$。[提示：考虑在练习 33 中所述的从(0, 0)到 $(n-k, k)$ 和从(0, 0)到 $(k, n-k)$ 的路径数。]

35. 使用练习 33 证明定理 4。[提示：计数练习 33 所描述的那种 n 步路径数。每条路径必须在一个 $(n-k, k)$ 点结束，其中 $k=0, 1, 2, \cdots, n$。]

36. 使用练习 33 证明帕斯卡恒等式。[提示：显示一条在练习 33 所描述的那种从(0, 0)到 $(n+1-k, k)$ 并通过 $(n+1-k, k-1)$ 点或 $(n-1, k)$ 点但不同时通过这两点的路径。]

37. 使用练习 33 证明练习 27 中的恒等式。[提示：首先观察到从(0, 0)到 $(n+1, r)$ 的路径数等于 $\binom{n+1+r}{r}$。其次，按照开始向上恰好走 k 个单位分别计数每一类路径，其中 $k=0, 1, 2, \cdots, r$，然后对结果求和。]

38. 如果 n 是正整数，则 $\sum_{k=0}^{n}k^2\binom{n}{k}=n(n+1)2^{n-2}$，给出一个组合证明。[提示：证明等式两边计数了从一个 n 元素集合中选一个子集，再从子集中选 2 个元素的方法，其中这 2 个元素可以相同。进一步，等式右边可以表示成 $n(n-1)2^{n-2}+n2^{n-1}$]

*39. 如果一个序列的前若干项如下列出，对于它的第 n 项确定一个与二项式系数有关的公式。[提示：对帕

斯卡三角形的观察有助于问题的求解。虽然以这一组给定的项作为开始的序列有无数多个，但下面列出的每个序列都是所求的那种序列的开始。]

a)1，3，6，10，15，21，28，36，45，55，66，…

b)1，4，10，20，35，56，84，120，165，220，…

c)1，2，6，20，70，252，924，3432，12 870，48 620，…

d)1，1，2，3，6，10，20，35，70，126，…

e)1，1，1，3，1，5，15，35，1，9，…

f)1，3，15，84，495，3003，18 564，116 280，735 471，4 686 825，…

4.5 一般性的排列和组合

4.5.1 引言



在许多计数问题里元素可以被重复使用。例如，一个字母或一个数字可以在一个车牌中多次使用。当选一打多纳圈时，每种可以被重复地选择。这与本章前面讨论的计数问题形成对照，因为在那里我们只考虑每个项至多可以使用一次的排列和组合。在这一节我们将介绍怎样求解元素可以多次使用的计数问题。

还有，某些计数问题涉及到不可区别的元素。例如，为计数单词 *SUCCESS* 的字母可能被重新排列的方式数，必须考虑相同字母的放置。这又与前面讨论的所有元素都被认为是不同的计数问题大相径庭。在这一节，我们将描述怎样求解某些元素是不可区别的计数问题。

此外，这一节也将解释怎样求解另一类重要的计数问题，即计数把不同的元素放入盒子的方法数的问题。这种问题的一个例子是把扑克牌发给 4 个玩牌人的不同的方式数。

把在本章前面描述的方法与这一节引入的方法一起考虑，就构成一个求解广泛的计数问题的有用的工具箱。当把第 5 章讨论的新方法再加到这个库里，你将能够求解在广泛的研究领域中产生的大多数计数问题。

4.5.2 有重复的排列

当元素允许重复时，使用乘积法则可以很容易地计数排列数，正如例 1 所示。

例 1 用英文字母可以构成多少个 n 位字符串？

解 因为有 26 个字母，且每个字母可以被重复使用，由乘积法则可以看出存在 26^n 个 n 位字符串。 ■

定理 1 给出了当允许重复时一个 n 元素集合的 r 排列数。

定理 1 具有 n 个物体的集合允许重复的 r 排列数是 n^r。

证 当允许重复时，在 r 排列中对 r 个位置中的每个位置有 n 种方式选择集合的元素，因为对每个选择，所有 n 个物体都是有效的。因此，由乘积法则，当允许重复时存在 n^r 个 r 排列。 □

4.5.3 有重复的组合

考虑下面元素允许重复的组合的实例。

例 2 从包含苹果、橙子和梨的碗里选 4 个水果。如果选择水果的顺序无关，且只关心水果的类型而不管是该类型的哪一个水果，那么当碗中每类水果至少有 4 个时有多少种选法？

解 为了求解这个问题，我们列出选择水果的所有可能的方式。共有 15 种方式：

4 个苹果	4 个橙子	4 个梨
3 个苹果，1 个橙子	3 个苹果，1 个梨	3 个橙子，1 个苹果
3 个橙子，1 个梨	3 个梨，1 个苹果	3 个梨，1 个橙子
2 个苹果，2 个橙子	2 个苹果，2 个梨	2 个橙子，2 个梨
2 个苹果，1 个橙子，1 个梨	2 个橙子，1 个苹果，1 个梨	2 个梨，1 个苹果，1 个橙子

这个解是从 3 个元素的集合{苹果，橙子，梨}中允许重复的 4 组合数。■

为求解这种类型的更复杂的计数问题，我们需要计数一个 n 元素集合的 r 组合的一般方法。在例 3 中，我们将给出这一方法。

例 3 从包含 1 美元、2 美元、5 美元、10 美元、20 美元、50 美元及 100 美元的钱袋中选 5 张纸币，有多少种方式？假定不管纸币被选的次序，同种币值的纸币都是不加区别的，并且至少每种纸币有 5 张。

解 因为纸币被选的次序是无关的且 7 种不同类型的纸币都可以选 5 次，问题涉及的是计数从 7 个元素的集合中允许重复的 5 组合数。列出所有的可能性将是乏味的，因为存在许多的解。相反，我们将给出一种方法来计数允许重复的组合数。

假设一个零钱盒子有 7 个隔间，每个保存一种纸币，如图 4-7 所示。这些隔间被 6 块隔板分开，正如图中所画的。每选择 1 张纸币就在相应的隔间里放置 1 个标记。图 4-8 针对选择 5 张纸币的 3 种不同方式给出了这种对应，其中的竖线表示 6 个隔板，星表示 5 张纸币。

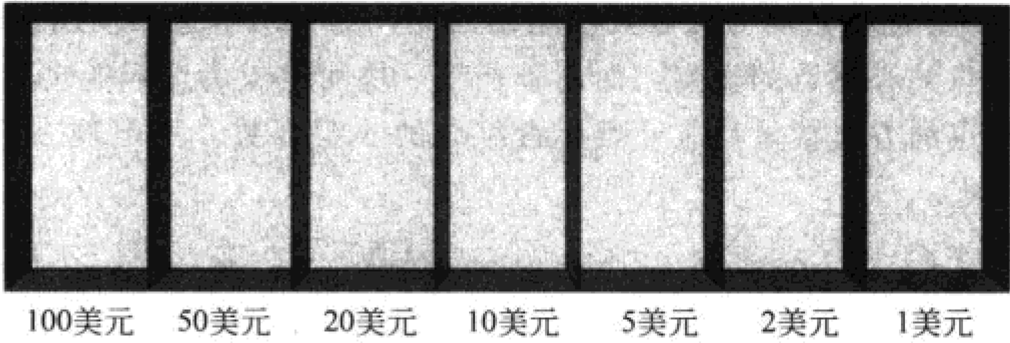


图 4-7　有 7 种类型纸币的零钱盒

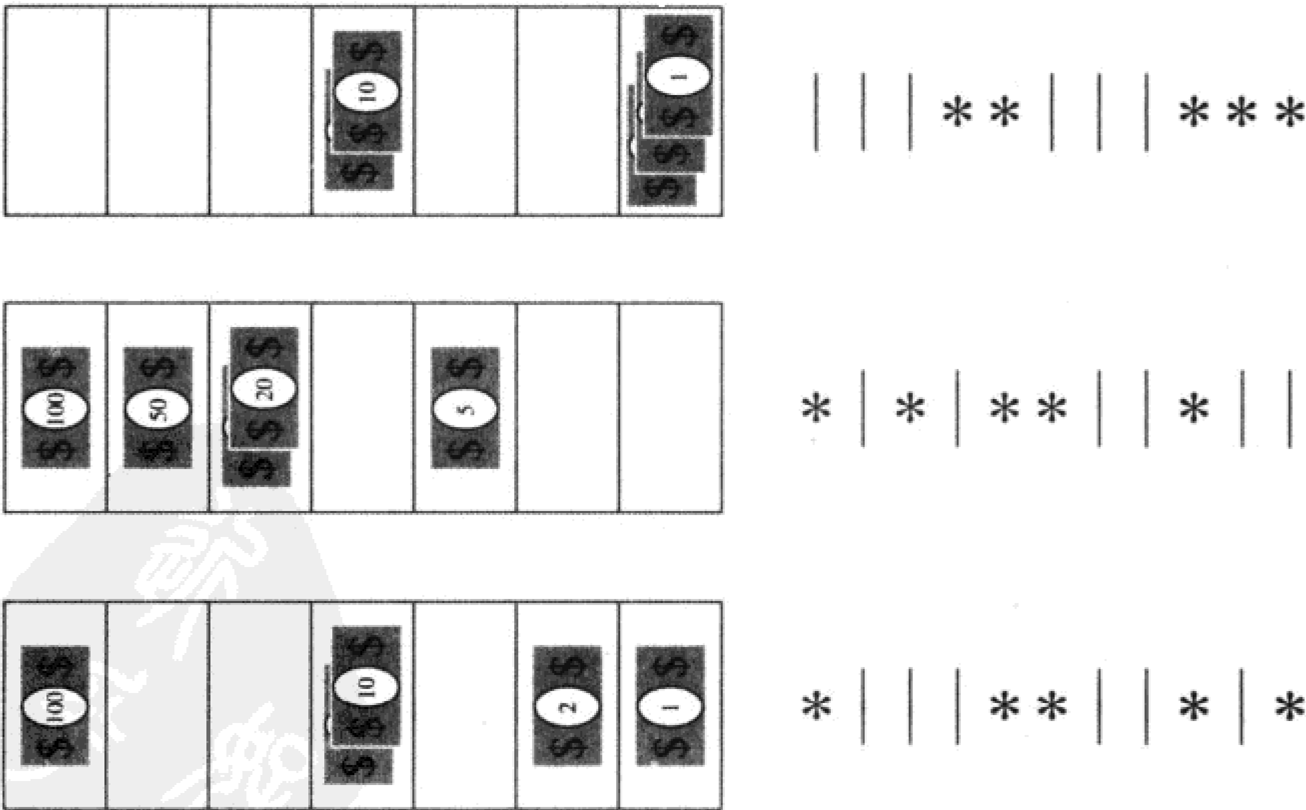


图 4-8　选择 5 张纸币的方式实例

选择 5 张纸币的方法数对应了安排 6 条竖线和 5 颗星的方法数。因此，选择 5 张纸币的方法数就是从 11 个可能的位置选 5 颗星位置的方法数。这对应了从含 11 个物体的集合中无序地选择 5 个物体的方法数，可以有 $C(11, 5)$ 种方式。因此存在

$$C(11,5)=\frac{11!}{5!6!}=462$$

种方式从有 7 类纸币的袋中选择 5 张纸币。 ■

定理 2 将这个讨论一般化。

定理 2 n 个元素的集合中允许重复的 r 组合有 $C(n+r-1, r)$ 个。

证 当允许重复时 n 元素集合的每个 r 组合可以用 $n-1$ 条竖线和 r 颗星的表表示。这 $n-1$ 条竖线用来标记 n 个不同的单元。每当集合的第 i 个元素出现在组合中，第 i 个单元就包含一颗星。例如，4 元素集合的一个 6 组合用 3 条竖线和 6 颗星来表示。这里

$$* * \mid * \mid\mid * * *$$

代表了恰包含 2 个第一元素、1 个第二元素、0 个第三元素和 3 个第四元素的组合。

正如我们已经看到的，包含 $n-1$ 条竖线和 r 颗星的每一个不同的表对应了 n 元素集合的允许重复的一个 r 组合。这种表的个数是 $C(n-1+r, r)$，因为每个表对应了从包含 r 颗星和 $n-1$ 条竖线的 $n-1+r$ 个位置中取 r 个位置来放 r 颗星的一种选择。 □

例 4 至例 7 说明怎样使用定理 2。

Extra Examples **例 4** 设一家甜点店有 4 种不同类型的甜点，那么从中选 6 块甜点有多少种不同的方式？假定只关心甜点的类型，而不管是哪一块甜点或者选择的次序。

解 选择 6 块甜点的方式数是具有 4 类元素集合的 6 组合数。由定理 2 这等于 $C(4+6-1, 6)=C(9, 6)$。由于

$$C(9,6)=C(9,3)=\frac{9\cdot 8\cdot 7}{1\cdot 2\cdot 3}=84$$

选择 6 块甜点的不同方式数有 84 种。 ■

定理 2 也可以用于求给定线性方程的整数解的个数。这可以由例 5 来说明。

例 5 方程

$$x_1+x_2+x_3=11$$

有多少个解？其中 x_1、x_2 和 x_3 是非负整数。

解 为计数解的个数，注意到一个解对应了从 3 元素集合中选 11 个元素的方式，以使得 x_1 选自第一类，x_2 选自第二类，x_3 选自第三类。因此，解的个数等于 3 元素集合允许重复的 11 组合数。由定理 2 存在

$$C(3+11-1,11)=C(13,11)=C(13,2)=\frac{13\cdot 12}{1\cdot 2}=78$$

个解。

当对变元加上限制时，也可以求出这个方程的解的个数。例如，当变元是满足 $x_1\geqslant 1$、$x_2\geqslant 2$ 且 $x_3\geqslant 3$ 的整数时，也可以求出这个方程的解的个数。满足此限制的方程的解对应于 11 个项的选择，使得项 x_1 取自第一类，项 x_2 取自第二类，项 x_3 取自第三类，并且第一类元素至少取 1 个，第二类元素至少取 2 个，第三类元素至少取 3 个。因此，先选 1 个第一类的元素，2 个第二类的元素，3 个第三类的元素；然后再多选 5 个元素。由定理 2，可以用

$$C(3+5-1,5)=C(7,5)=C(7,2)=\frac{7\cdot 6}{1\cdot 2}=21$$

种方式做到。于是，对给定限制的方程存在 21 个解。 ■

在一个 n 元素集合允许重复的 r 组合和把 r 个相同的球放到 n 个不同的箱子的方法之间存在一一对应。为建立这种对应，每次我们把一个球放到第 i 个箱子里，集合的第 i 个元素就加到这个 r 组合中。

例 6　有多少种方法把 10 个相同的球放人 8 个不同的箱子？

解　把 10 个相同的球放入 8 个不同箱子的方法数等于 8 个元素的集合允许重复的 10 组合数，因此存在

$$C(8+10-1,10)=C(17,10)=\frac{17!}{10!7!}=19\,448$$

种方法。■

例 7 显示了怎样计数在确定变量值时产生的允许重复的组合数，当每次通过某一类确定的嵌套循环时，这个变量的值都会增加。

例 7　在下面的伪码被执行后 k 的值是什么？

```
k := 0
for i1 := 1 to n
    for i2 := 1 to i1
        ⋮
    for im := 1 to im-1
        k := k+1
```

解　k 的初值是 0，且对于满足

$$1\leqslant i_m\leqslant i_{m-1}\leqslant\cdots\leqslant i_1\leqslant n$$

的整数序列 i_1，i_2，…，i_m，每次通过这个嵌套循环时 k 的值就加 1。这种整数序列的个数是从 $\{1,2,\cdots,n\}$ 中允许重复地选择 m 个整数的方式数。（为看到这一点只需注意到一旦这个整数序列选定以后，如果我们按非降顺序排列序列中的整数，这就唯一地确定了一组对 i_m，i_{m-1}，…，i_1 的赋值。相反，每个这样的赋值对应了一个唯一的无序集合。）所以由定理 2 得出在这个代码被执行后 $k=C(n+m-1,m)$。■

从一个 n 元素集合中，允许重复和不允许重复地选择 r 个元素，其有序或无序的选择数的公式在表 4-1 给出。

表 4-1　允许和不允许重复的组合与排列

类　型	允许重复	公　　式
r 排列	不	$\frac{n!}{(n-r)!}$
r 组合	不	$\frac{n!}{r!\,(n-r)!}$
r 排列	是	n^r
r 组合	是	$\frac{(n+r-1)!}{r!\,(n-1)!}$

4.5.4　具有不可区别物体的集合的排列

在计数问题中某些元素可能是没有区别的。在这种情况下必须小心避免重复计数。考虑例 8。

Extra Examples　**例 8**　重新排序单词 *SUCCESS* 中的字母能构成多少个不同的串？

解　因为 *SUCCESS* 中的某些字母是重复的，答案并不是 7 个字母的排列数。这个单词包含 3 个 S、2 个 C、1 个 U 和 1 个 E。为确定重新排序单词中的字母能构成多少个不同的串，

首先，注意到3个S可以用$C(7,3)$种不同的方式放在7个位置中，剩下4个空位。然后可以用$C(4,2)$种方式放2个C，留下2个空位。又可以用$C(2,1)$种方式放U，留下1个空位。因此，放E只有$C(1,1)$种方式。从而，由乘积法则，产生的不同的串数是

$$\begin{aligned}C(7,3)C(4,2)C(2,1)C(1,1) &= \frac{7!}{3!4!}\cdot\frac{4!}{2!2!}\cdot\frac{2!}{1!1!}\cdot\frac{1!}{1!0!}\\ &= \frac{7!}{3!2!1!1!}\\ &= 420\end{aligned}$$

■

使用和例8同样的推理，能够证明定理3。

定理3 设类型1的相同的物体有n_1个，类型2的相同的物体有n_2个，…，类型k的相同的物体有n_k个，那么n个物体的不同排列数是

$$\frac{n!}{n_1!n_2!\cdots n_k!}$$

证 为确定排列数，首先注意到可以用$C(n,n_1)$种方式在n个位置中放类型1的n_1个物体，剩下$n-n_1$个空位。然后用$C(n-n_1,n_2)$种方式放类型2的物体，剩下$n-n_1-n_2$个空位。继续放类型3的物体，…，类型$k-1$的物体，直到最后可用$C(n-n_1-n_2-\cdots-n_{k-1},n_k)$种方式放类型$k$的物体。因此由乘积法则，不同排列的总数是

$$\begin{aligned}&C(n,n_1)C(n-n_1,n_2)\cdots C(n-n_1-\cdots-n_{k-1},n_k)\\ &= \frac{n!}{n_1!(n-n_1)!}\cdot\frac{(n-n_1)!}{n_2!(n-n_1-n_2)!}\cdot\cdots\cdot\frac{(n-n_1-\cdots-n_{k-1})!}{n_k!0!}\\ &= \frac{n!}{n_1!n_2!\cdots n_k!}\end{aligned}$$

□

4.5.5 把物体放入盒子

有些计数问题可以通过枚举把不同的物体放入不同的盒子的方式数来求解。考虑下面的例子，其中物体是牌且“盒子”是玩牌人的手。

例9 有多少种方式把52张标准的扑克牌发给4个人使得每个人5张牌？

解 我们将使用乘积法则求解这个问题。开始，第一个人得到5张牌可以有$C(52,5)$种方式。第二个人得到5张牌可以有$C(47,5)$种方式，因为只剩下47张牌。第三个人得到5张牌可以有$C(42,5)$种方式。最后，第四个人得到5张牌可以有$C(37,5)$种方式。因此，发给4个人每人5张牌的方式总数是

$$\begin{aligned}C(52,5)C(47,5)C(42,5)C(37,5) &= \frac{52!}{47!5!}\cdot\frac{47!}{42!5!}\cdot\frac{42!}{37!5!}\cdot\frac{37!}{32!5!}\\ &= \frac{52!}{5!5!5!5!32!}\end{aligned}$$

■

注意 例9的解等于52个物体的排列数，这些物体分成5个不同的类，其中四类，每类有5个相同的物体，第五类有32个物体。可以通过在这种排列和给人发牌之间定义一个一一对应来说明这个等式。为定义这个对应，首先把牌从1到52排序。然后将发给第一个人的牌与分配给第一类物体在排列中的位置对应。类似地，发给第二、第三和第四个人的牌分别与第二、第三、第四类物体所分配的位置对应。没有发给任何人的牌与第五类物体所分配的位置对应。读者应该能够验证这是一个一一对应。

例9是涉及把不同的物体分配到不同的盒子的一个典型的问题。这些不同的物体是52张牌，

5 个不同的盒子是 4 个人的手和其余的牌。可以使用下面的定理求解把不同的物体分配到不同的盒子的计数问题。

定理 4 把 n 个不同的物体分配到 k 个不同的盒子使得 n_i 个物体放入盒子 $i(i=1, 2, \cdots, k)$ 的方式数等于

$$\frac{n!}{n_1!n_2!\cdots n_k!}$$

定理 4 可以使用乘积法则证明。详细证明见本节的练习 47。它也可以通过在定理 3 计数的排列和定理 4 计数的放物体的方法之间建立一一对应来给出证明(见练习 48)。

练习

1. 从一个 3 元素集合中允许重复地有序选取 5 个元素有多少种不同的方式?
2. 从一个 5 元素集合中允许重复地有序选取 5 个元素有多少种不同的方式?
3. 6 个字母的字符串有多少个?
4. 每天一个学生从一堆包好的三明治中随机选 1 块三明治作为午饭。如果有 6 种三明治并且选择三明治的次序无关,在一周的 7 天里这个学生选择三明治有多少种不同的方式?
5. 分配 3 种工作给 5 个雇员,如果每个雇员可以得到 1 种以上的工作,那么有多少种不同的分配方式?
6. 从一个 3 元素集合中允许重复地无序选取 5 个元素有多少种不同的方式?
7. 从一个 5 元素集合中允许重复地无序选取 3 个元素有多少种不同的方式?
8. 从一个商店的 21 种多纳圈中选择 12 个多纳圈有多少种不同的方式?
9. 一个百吉饼店有洋葱百吉饼、罂粟子百吉饼、鸡蛋百吉饼、咸味百吉饼、粗制倮麦百吉饼、芝麻百吉饼、葡萄干百吉饼和普通百吉饼,有多少种方式选择
   a)6 个百吉饼?
   b)12 个百吉饼?
   c)24 个百吉饼?
   d)12 个百吉饼,并且每类至少有 1 个?
   e)12 个百吉饼,并且至少有 3 个鸡蛋百吉饼和不超过 2 个咸味百吉饼?
10. 一个新月形面包店有普通新月形面包、樱桃新月形面包、巧克力新月形面包、杏仁新月形面包、苹果新月形面包和椰菜新月形面包。有多少种方式选择
    a)12 个新月形面包?
    b)36 个新月形面包?
    c)24 个新月形面包,并且至少每类有 2 个?
    d)24 个新月形面包,并且不超过 2 个椰菜的?
    e)24 个新月形面包,并且至少 5 个巧克力的且至少 3 个杏仁的?
    f)24 个新月形面包,并且至少 1 个普通的,至少 2 个樱桃的,至少 3 个巧克力的,至少 1 个杏仁的,至少 2 个苹果的和不超过 3 个椰菜的?
11. 一个小猪储钱罐包含 100 个相同的 1 美分和 80 个相同的 5 美分硬币,从中选 8 个硬币有多少种方式?
12. 如果一个小猪储钱罐中有 1 美分、5 美分、10 美分、25 美分、50 美分等硬币,那么 20 个硬币有多少种不同的组合?
13. 一个出版商有 3000 本离散数学书,如果这些书是没有区别的,那么将这些书存储在 3 个库房有多少种方式?
14. 设 x_1、x_2、x_3 和 x_4 是非负整数,方程
    $$x_1+x_2+x_3+x_4=17$$
    有多少个解?
15. 方程

$$x_1 + x_2 + x_3 + x_4 + x_5 = 21$$

有多少个解？其中 $x_i(i=1, 2, 3, 4, 5)$是非负整数，并且使得

a) $x_1 \geqslant 1$

b) $x_i \geqslant 2$，$i=1, 2, 3, 4, 5$

c) $0 \leqslant x_1 \leqslant 10$

d) $0 \leqslant x_1 \leqslant 3$，$1 \leqslant x_2 < 4$，$x_3 \geqslant 15$

16. 方程

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 29$$

有多少个解？其中 $x_i(i=1, 2, 3, 4, 5, 6)$是非负整数，并且使得

a) $x_i \geqslant 1$，$i=1, 2, 3, 4, 5, 6$

b) $x_1 \geqslant 1$，$x_2 \geqslant 2$，$x_3 \geqslant 3$，$x_4 \geqslant 4$，$x_5 > 5$，$x_6 \geqslant 6$

c) $x_1 \leqslant 5$

d) $x_1 < 8$，$x_2 > 8$

17. 有多少 10 位三进制数字(0，1 或 2)串恰含有 2 个 0、3 个 1 和 5 个 2?

18. 有多少 20 位十进制数字串含有 2 个 0、4 个 1、3 个 2、1 个 3、2 个 4、3 个 5、2 个 7 和 3 个 9?

19. 假设一个大家庭有 14 个孩子，包括 2 组三胞胎、3 组双胞胎以及 2 个单胞胎。这些孩子坐在一排椅子上，如果相同的三胞胎或双胞胎的孩子不能互相区分，那么有多少种方式?

20. 不等式

$$x_1 + x_2 + x_3 \leqslant 11$$

有多少个解？其中 x_1、x_2 和 x_3 是非负整数[提示：引入辅助变量 x_4 使得 $x_1+x_2+x_3+x_4=11$。]

21. 把 6 个相同的球放到 9 个不同的箱子里有多少种方法?

22. 把 12 个相同的球放到 6 个不同的箱子里有多少种方法?

23. 把 12 个不同的物体放到 6 个不同的盒子里并且每个盒子有 2 个物体，有多少种方法?

24. 把 15 个不同的物体放到 5 个不同的盒子里并且这些盒子分别有 1 个、2 个、3 个、4 个和 5 个物体，有多少种方法?

25. 有多少个小于 1 000 000 的正整数其数字之和等于 19?

26. 有多少个小于 1 000 000 的正整数恰好一个数字等于 9 且数字之和等于 13?

27. 一次离散数学的期终考试有 10 道题。如果总分数是 100 且每道题至少 5 分，那么有多少种方式来分配这些题的分数?

28. n 个物体有 r 种不同的类型，证明有 $C(n+r-q_1-q_2-\cdots-q_r-1, n-q_1-q_2-\cdots-q_r)$种不同的无序选择，使得该选择至少有 q_1 个 1 型的物体，q_2 个 2 型物体，…，q_r 个 r 型物体?

29. 如果被传送的二进制串必须以 1 开始，必须有另外 3 位 1(使得传送的 1 共有 4 位)，必须包含总共 12 位 0，必须每个 1 后面至少跟随 2 个 0，那么有多少个不同的二进制串?

30. 使用 *MISSISSIPPI* 中的所有字母可以构造多少个不同的串?

31. 使用 *ABRACADABRA* 中的所有字母可以构造多少个不同的串?

32. 使用 *AARDVARK* 中的所有字母且所有的 3 个 *A* 必须连续，可以构造多少个不同的串?

33. 使用 *ORONO* 中的某些或全部字母可以构造多少个不同的串?

34. 使用 *SEERESS* 中的字母可以构造多少个至少含 5 个字符的串?

35. 用 *EVERGREEN* 中的字母可以构造多少个至少含 7 个字符的串?

36. 使用 6 个 1 和 8 个 0 可以构造多少个不同的二进制串?

37. 一个学生有 3 个芒果、2 个番木瓜和 2 个猕猴桃。如果这个学生每天吃 1 个水果，并且只考虑水果的类型，那么有多少种不同的方式吃完这些水果?

38. 一个教授把 40 本数学期刊放入 4 个盒子，每盒 10 本，分配这些期刊有多少种方式?

a)如果每个盒子被编号使得它们是可区分的。

b)这些盒子是相同的，使得它们是不可区分的。

39. 有多少种不同的方式在 xyz 空间上从原点(0，0，0)到达(4，3，5)点？这个旅行的每一步是在 x 正方向移动一个单位，y 正方向移动一个单位，或者 z 正方向移动一个单位。(x，y，z 负方向的移动是禁止的，即不允许回头。)
40. 有多少种不同的方式在 $xyzw$ 空间上从原点(0，0，0，0)到达(4，3，5，4)点？这个旅行的每一步是在 x、y、z 或 w 正方向移动一个单位。
41. 把一副标准的 52 张扑克牌发给 5 个人，每个人得到 7 张牌，有多少种方式？
42. 在打桥牌时，把一副标准的 52 张牌发给 4 个人，有多少种不同发牌的方式？
43. 当把一副标准的 52 张牌发给 4 个人时，若使得每个人有一手包含 1 个 A 的牌，这种概率是多少？
44. 12 本书放在 4 个不同的书架上有多少种方式？
    a)如果这些书是同一种书。
    b)如果所有的书都不同，并且考虑这些书在书架上的位置。[提示：把这件事分成 12 个任务完成，放每本书是一个任务。先用 1，2，3，4 表示这些书架，用 b_i($i=1$，2，…，12)表示书。把 b_i 放到 1，2，3，4 中某个数的右边。]
45. n 本书放在 k 个不同的书架上有多少种方式？
    a)如果这些书是同一种书。
    b)如果所有的书都不同，并且考虑这些书在书架上的位置。
46. 12 本书在一个书架上排成一排。从中选 5 本书并且使得没有 2 本书相邻有多少种方式？[提示：将选的书用竖线表示，没选的书用星表示，计数含 5 条竖线和 7 颗星且没有 2 条竖线相邻的序列数。]

*47. 通过先把物体放入第一个盒子，然后把物体放入第二个盒子，…，的方法，使用乘积法则证明定理 4。

*48. 通过下面的方法证明定理 4。有 n 个物体，其中类型为 i 的相同的物体有 n_i 个，$i=1$，2，…，k。先把这 n 个物体的排列和把这些物体放到 k 个盒子且使得盒子 i 含有 n_i 个物体的分配之间建立一一对应，这里 $i=1$，2，…，k，然后使用定理 3。

*49. 在这个练习中我们将通过在两个集合之间建立一一对应来证明定理 2。这两个集合分别是集合 $S=\{1,2,\cdots,n\}$ 的允许重复的 r 组合的集合和集合 $T=\{1,2,3,\cdots,n+r-1\}$ 的 r 组合的集合。
    a)把 S 的允许重复的 r 组合中的元素排成一个递增序列 $x_1\leqslant x_2\leqslant\cdots\leqslant x_r$。证明对这个序列的第 k 项加上 $k-1$ 而构成的序列是严格递增的。断言这个序列由 T 的 r 个不同的元素构成。
    b)证明 a)所描述的过程在 S 的允许重复的 r 组合的集合与 T 的 r 组合的集合之间定义了一一对应。[提示：通过把 T 的满足 $1\leqslant x_1<x_2<\cdots<x_r\leqslant n+r-1$ 的 r 组合 $\{x_1,x_2,\cdots,x_r\}$，与从第 k 个元素减去 $k-1$ 得到的 S 的允许重复的 r 组合相联系，证明这个对应是可逆的。]
    c)断言存在着 $C(n+r-1,r)$ 个 n 元素集合的允许重复的 r 组合。

50. 有多少种方式把 5 个不同的物体放到 3 个相同的盒子里？
51. 有多少种方式把 5 个相同的物体放到 3 个相同的盒子里？
52. 在 $(x_1+x_2+\cdots+x_m)^n$ 的展开式中，把所有的同类项合并以后有多少个不同的项？

*53. 证明多项式定理：如果 n 是正整数，则

$$(x_1+x_2+\cdots+x_m)^n = \sum_{n_1+n_2+\cdots n_m=n} C(n;n_1,n_2,\cdots,n_m)x_1^{n_1}x_2^{n_2}\cdots x_n^{n_m},$$

其中

$$C(n;n_1,n_2,\cdots,n_m)=\frac{n!}{n_1!n_2!\cdots n_m!}$$

是多项式系数。

54. 求 $(x+y+z)^4$ 的展开式。
55. 求 $(x+y+z)^{10}$ 中的 $x^3y^2z^5$ 的系数。
56. 在 $(x+y+z)^{100}$ 的展开式中有多少个项？

4.6 生成排列和组合

4.6.1 引言

本章前几节已经描述了各种类型的排列和组合的计数方法，但是有时候需要生成排列和组合，而不仅仅是计数。考虑下面三个问题。第一，假设一个销售商必须访问 6 个城市。应该按照什么顺序访问这些城市而使得总的旅行时间最少？一种方法就是确定 6！＝720 种不同顺序的访问时间并且选择具有最小旅行时间的访问顺序。第二，假定 6 个数的集合中某些数的和是 100。找出这些数的一种方法就是生成所有 $2^6=64$ 个子集并且检查它们的元素和。第三，假设一个实验室有 95 个雇员，一个项目需要一组 12 人组成的有 25 种特定技能的雇员。（每个雇员可能有一种或多种技能）。找出这组雇员的一种方法就是找出所有的 12 个雇员的小组，然后检查他们是否有所需要的技能。这些例子都说明为了求解问题常常需要生成排列和组合。

4.6.2 生成排列

任何 n 元素集合可以与集合$\{1, 2, 3, \cdots, n\}$建立一一对应。我们可以如下列出任何 n 元素集合的所有排列：生成 n 个最小正整数的排列，然后用对应的元素替换这些整数。已经建立了许多不同的算法来生成这个集合的 $n!$ 个排列。我们将要描述的算法是以$\{1, 2, \cdots, n\}$的排列集合上的*字典顺序*为基础的。按照这个顺序，如果对于某个 k，$1\leqslant k\leqslant n$，$a_1=b_1$，$a_2=b_2$，…，$a_{k-1}=b_{k-1}$并且 $a_k<b_k$，那么排列 $a_1a_2\cdots a_n$ 在排列 $b_1b_2\cdots b_n$ 的前边。换句话说，如果在 n 个最小正整数集合的两个排列不等的第一位置，一个排列的数小于第二个排列的数，那么这个排列按照字典顺序排在第二个排列的前边。

例 1 集合$\{1, 2, 3, 4, 5\}$的排列 23415 在排列 23514 的前边，因为这些排列在前两位相同，但第一排列在第三位置中的数是 4，小于第二排列在第三位置中的数 5。类似地，排列 41532 在排列 52143 的前边。■

生成$\{1, 2, \cdots, n\}$的排列的算法基础是从一个给定排列 $a_1a_2\cdots a_n$ 按照字典顺序构造下一个排列的过程。我们将说明怎样做到这一点。首先假设 $a_{n-1}<a_n$，交换 a_{n-1}和 a_n 可得到一个更大的排列。没有其他的排列既大于原来的排列并且又小于这个通过交换 a_{n-1}与 a_n 得到的排列。例如，在 234156 后面的下一个最大的排列是 234165。另一方面，如果 $a_{n-1}>a_n$，那么由交换这个排列中的最后两项不可能得到一个更大的排列。看看排列中的最后 3 个整数，如果 $a_{n-2}<a_{n-1}$，那么可以重新安排这后 3 个数而得到下一个最大的排列。a_{n-1}和 a_n 中比较小的数大于 a_{n-2}，先把这个数放在位置 $n-2$，然后把剩下的那个数和 a_{n-2}按照递增的顺序放到最后的两个位置。例如，在 234165 后面的下一个最大的排列是 234516。

另一方面，如果 $a_{n-2}>a_{n-1}$（且 $a_{n-1}>a_n$），那么不可能由安排在这个排列的最后三项而得到更大的排列。基于这个观察，可以描述一个一般的方法，对于给定的排列 $a_1a_2\cdots a_n$ 依据字典顺序来生成下一个最大的排列。首先，找到整数 a_j 和 a_{j+1}使得 $a_j<a_{j+1}$且

$$a_{j+1}>a_{j+2}>\cdots>a_n,$$

即在这个排列中的最后一对相邻的整数，使得这个对的第一个整数小于第二个整数。然后，把 a_{j+1}，a_{j+2}，…，a_n 中大于 a_j 的最小的整数放到第 j 个位置，再按照递增顺序从位置 $j+1$ 到 n 列出 a_j，a_{j+1}，a_{j+2}，…，a_n 中其余的整数，这就得到依照字典顺序的下一个最大的排列。容易看出，没有其他的排列大于排列 $a_1a_2\cdots a_n$ 而小于这个新生成的排列。（对这一事实的验证留给读者作为练习）。

例 2　在 362541 后面按照字典顺序下一个最大的排列是什么？

解　使得 $a_j < a_{j+1}$ 的最后一对整数 a_j 和 a_{j+1} 是 $a_3=2$ 和 $a_4=5$。排列在 2 右边大于 2 的最小整数是 $a_5=4$，因此 4 被放在第三位置。然后整数 2，5 和 1 依递增顺序放到最后 3 个位置，即这个排列的最后 3 个位置是 125。于是，下一个最大排列是 364125。■

为生成整数 1，2，…，n 的 $n!$ 个排列，按照字典顺序由最小的排列，即 $123\cdots n$ 开始，连续施用 $n!-1$ 次生成下一个最大排列的过程，就得到 n 个最小的整数按字典顺序的所有排列。

例 3　按字典顺序生成整数 1，2，3 的排列。

解　从 123 开始，由交换 3 和 2 得到下一个排列 132。下一步，因为 $3>2$ 和 $1<3$，排列在 132 中的 3 个整数，把 3 和 2 中较小的放到第一个位置，然后按递增顺序把 1 和 3 放到位置 2 和 3 而得到 213。跟着 213 的是 231，它是由交换 1 和 3 得到的，因为 $1<3$。下一个最大的排列把 3 放在第一位置，后面是 1 和 2 按递增顺序排列，即 312。最后，交换 1 和 2 得到最后一个排列 321。■

算法 1 显示了在给定排列不是最大的排列 $n\ n-1\ n-2\cdots 2\ 1$ 时，在它的后面按照字典顺序找到下一个最大排列的过程。

算法 1　按字典顺序生成下一个最大排列

```
procedure next permutation(a1 a2 … an: {1, 2, …, n}的排列，不等于 n n−1 …2 1)
j := n−1
while aj > aj+1
  j := j−1
{j 是使得 aj < aj+1 的最大下标}
k := n
while aj > ak
  k := k−1
{ak 是在 aj 右边大于 aj 的最小整数}
交换 aj 和 ak
r := n
s := j+1
while r > s
begin
  交换 ar 和 as
  r := r−1
  s := s+1
end
{这把在第 j 位后边的排列尾部按递增顺序放置}
```

4.6.3 生成组合

怎样可以生成一个有穷集的元素的所有组合呢？由于一个组合仅仅就是一个子集，我们可以利用在 $\{a_1, a_2, \cdots, a_n\}$ 和 n 位二进制串之间的对应。

如果 a_k 在子集中，对应的二进制串在位置 k 有一个 1；如果 a_k 不在子集中，对应的二进制串在位置 k 有一个 0。如果可以列出所有的 n 位二进制串，那么通过在子集和二进制串之间的对

应就可以列出所有的子集。

一个 n 位二进位串也是一个在 0 到 2^n-1 之间的整数的二进制展开式。按照它们的二进制展开式，作为整数根据递增顺序可以列出这 2^n-1 个二进制串。为生成所有的 n 位二进制展开式，从具有 n 个 0 的二进位串 000…00 开始。然后，继续找下一个最大的展开式，直到得到 111…11 为止。在每一步找下一个最大的二进制展开式时，先确定从右边起第一个不是 1 的位置。然后把这个位置右边的所有的 1 变成 0 并且将这第一个 0(从右边数)变成 1。

例 4 找出在 10 0010 0111 后面的下一个最大的二进位串。

解 这个串从右边数不是 1 的第 1 位是从右边起的第 4 位。把这一位变成 1 并且将它后面所有的位变成 0。这就生成了下一个最大的二进位串 10 0010 1000。 ■

生成在 $b_{n-1}b_{n-2}\cdots b_1b_0$ 后面的下一个最大的二进位串的过程给出在算法 2 中。

算法 2 生成下一个最大的二进位串

```
procedure next bit string(b_{n-1}b_{n-2}…b_1b_0: 不等于 11…11 的二进位串)
i := 0
while b_i = 1
begin
  b_i := 0
  i := i+1
end
  b_i := 1
```

下面将给出生成集合{1, 2, 3, …, n}的 r 组合的算法。一个 r 组合可以表示成一个序列，这个序列按照递增的顺序包含了这个子集中的元素。使用在这些序列上的字典顺序可以列出这些 r 组合。在 $a_1a_2\cdots a_r$ 后面的下一个组合可以按下面的方法得到：首先，找到序列中使得 $a_i\neq n-r+i$ 的最后元素 a_i，然后用 a_i+1 代替 a_i，且对于 $j=i+1$，$i+2$，…，r 用 $a_i+j-i+1$ 代替 a_j。请读者证明这就按字典顺序生成了下一个最大的组合。下面的例子说明了这个过程。

例 5 找出集合{1, 2, 3, 4, 5, 6}在{1, 2, 5, 6}后面的下一个最大的 4 组合。

解 在具有 $a_1=1$，$a_2=2$，$a_3=5$，$a_4=6$ 的项中使得 $a_i\neq 6-4+i$ 的最后的项是 $a_2=2$。为得到下一个最大的 4 组合，把 a_2 加 1 得 $a_2=3$。然后，置 $a_3=3+1=4$ 且 $a_4=3+2=5$。从而下一个最大的 4 组合是{1, 3, 4, 5}。 ■

算法 3 用伪码给出了这个过程。

算法 3 按字典顺序生成下一个 r 组合

```
procedure next r-combination({a_1, a_2, …, a_r}: {1, 2, …, n}的满足 a_1<a_2<…<a_r 的不等于{n-r+1, …, n}的真子集)
i := r
while a_i = n-r+i
  i := i-1
a_i := a_i+1
for j := i+1 to r
  a_j := a_i+j-i
```

练习

1. 找出按照字典顺序跟在下面每一个排列后面的下一个最大的排列。
   a)1432　　b)54123　　c)12453
   d)45231　　e)6714235　　f)31528764
2. 找出按照字典顺序跟在下面每一个排列后面的下一个最大的排列。
   a)1342　　b)45321　　c)13245
   d)612345　　e)1623547　　f)23587416
3. 按照字典顺序排列下述{1，2，3，4，5}的排列：43521，15432，45321，23451，23514，14532，21345，45213，31452，31542。
4. 按照字典顺序排列下述{1，2，3，4，5，6}的排列：234561，231456，165432，156423，543216，541236，231465，314562，432561，654321，654312，435612。
5. 使用算法 1 按照字典顺序生成前 4 个正整数的 24 个排列。
6. 使用算法 2 列出集合{1，2，3，4}的所有子集。
7. 使用算法 3 列出集合{1，2，3，4，5}的所有的 3 组合。
8. 证明算法 1 按字典顺序生成下一个最大的排列。
9. 证明算法 3 按字典顺序生成给定 r 组合后面的下一个最大的 r 组合。
10. 建立一个算法来生成 n 元素集合的 r 排列。
11. 列出{1，2，3，4，5}的所有 3 排列。

这一节剩下的练习建立了另一个算法来生成{1，2，3，…，n}的排列。这个算法是基于整数的康托尔展开。每个小于 $n!$ 的非负整数有一个唯一的康托尔展开式

$$a_1 1! + a_2 2! + \cdots + a_{n-1}(n-1)!$$

其中 a_i 是一个不超过 i 的非负整数，$i=1，2，\cdots，n-1$。整数 $a_1，a_2，\cdots，a_{n-1}$ 叫做这个整数的**康托尔数字**。

给定{1，2，…，n}的一个排列。令 a_{k-1} 是排列中在 k 后面且小于 k 的整数个数，$k=2，3，\cdots，n$。例如，在排列 43215 中，a_1 是在 2 后面且小于 2 的整数个数，所以 $a_1=1$。类似地，对这个例子 $a_2=2$，$a_3=3$ 且 $a_4=0$。考虑从{1，2，3，…，n}的排列的集合到小于 $n!$ 的非负整数的集合的函数。这个函数把一个排列映到一个非负整数，而这个整数把以这种方式定义的 $a_1，a_2，\cdots，a_{n-1}$ 作为它的康托尔数字。

12. 找出对应于下述排列的整数。
    a)246531　　b)12345　　c)654321
*13. 证明这里描述的对应是{1，2，3，…，n}的排列的集合与小于 $n!$ 的非负整数之间的双射。
14. 按照康托尔展开式与练习 12 前面所描述的排列之间的对应找出与下面的整数相对应的{1，2，3，4，5}的排列。
    a)3　　b)89　　c)111
15. 开发一个以练习 12 描述的对应为基础的算法来生成 n 元素集合所有的排列。
*16. 下面的方法可以用来生成一个 n 项序列的随机排列。首先，交换第 n 项与第 $r(n)$ 项，其中 $r(n)$ 是一个满足 $1 \leqslant r(n) \leqslant n$ 的随机选择的整数。接着，交换结果序列的第$(n-1)$项与第 $r(n-1)$ 项，其中 $r(n-1)$ 是一个满足 $1 \leqslant r(n-1) \leqslant n-1$ 的随机选择的整数。继续这一过程直到 $j=n$，其中在第 j 步要交换结果序列的第$(n-j+1)$项与第 $r(n-j+1)$ 项，其中 $r(n-j+1)$ 是一个满足 $1 \leqslant r(n-j+1) \leqslant n-j+1$ 的随机选择的整数。证明当遵循这个方法时序列项的 $n!$ 个不同排列中的每一个排列被等可能地生成。[提示：使用数学归纳法。对于 $n-1$ 项的一个序列，假定这一过程生成的每一个 $n-1$ 项的排列的概率是 $1/(n-1)!$。]

关键术语和结果

术语

组合数学：研究物体安排的科学。

枚举：物体安排的计数。

树图：由根、从根出发的分支以及从分支的某些端点出发的其他分支构成的图。

排列：集合元素的一个有序的安排。

r 排列：集合的 r 个元素的一个有序安排。

$P(n, r)$：n 元素集合的 r 排列数。

r 组合：集合的 r 个元素的无序选取。

$C(n, r)$：n 元素集合的 r 组合数。

$\binom{n}{r}$（二项式系数）：也是 n 元素集合的 r 组合数。

组合证明：基于计数变量的证明。

帕斯卡三角形：二项式系数的一种表示，其中三角形的第 i 行包含 $\binom{i}{j}$，$j=0, 1, 2, \cdots, i$。

结果

乘积法则：一种基本的计数技术。这个法则指出，当一个过程由两个子任务构成时，完成这个过程的方式数是完成第一个任务的方式数和完成第一个任务之后再做第二个任务的方式数之积。

求和法则：一种基本的计数技术。这个法则指出，如果两个任务不能同时做，那么用这种或那种方式完成任务的总方式数是完成两种任务的方式数之和。

鸽巢原理：当比 k 多的物体放到 k 个盒子时，一定存在一个盒子包含了至少 2 个物体。

广义鸽巢原理：当 N 个物体放入 k 个盒子时，一定存在一个盒子包含了至少 $\lceil N/k \rceil$ 个物体。

$$P(n,r) = \frac{n!}{(n-r)!}$$

$$C(n,r) = \binom{n}{r} = \frac{n!}{r!(n-r)!}$$

帕斯卡恒等式：$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$

二项式定理：$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^k$

当允许重复时，一个 n 元素集合有 n^r 个 r 排列。

当允许重复时，一个 n 元素集合有 $C(n+r-1, r)$ 个 r 组合。

如果类型 i 的不可区分的物体有 n_i 个，$i=1, 2, 3, \cdots, k$，那么 n 个物体的排列有 $n! / (n_1! n_2! \cdots n_k!)$ 个。

生成集合 $\{1, 2, \cdots, n\}$ 的排列的算法。

复习题

1. 解释怎样用求和与乘积法则找出长度不超过 10 的二进位串的个数。
2. 解释怎样找出长度不超过 10 且至少有 1 位 0 的二进位串的个数。
3. a)怎样用乘积法则找出从 m 元素集合到 n 元素集合的函数个数？
   b)从一个 5 元素集合到一个 10 元素集合存在多少个函数？
   c)怎样用乘积法则找出从 m 元素集合到 n 元素集合的一对一函数的个数？
   d)从一个 5 元素集合到一个 10 元素集合存在多少个一对一函数？
   e)从一个 5 元素集合到一个 10 元素集合存在多少个映上（注：满射）的函数？
4. 如果首先赢 4 个球的队就能取胜，你怎样找出两个队加赛的所有可能的结果数？
5. 你怎样找出以 101 开始或以 010 结束的 10 位二进位串数？

6. a)叙述鸽巢原理。
   b)解释怎样用鸽巢原理证明在 11 个整数中至少 2 个整数的最后一位相同?
7. a)叙述广义鸽巢原理。
   b)解释怎样用广义鸽巢原理证明在 91 个整数中有 10 个整数的最后一位数字相同?
8. a)一个 n 元素集合的 r 排列和 r 组合的区别是什么?
   b)推导一个与 n 元素集合的 r 组合数及 r 排列数有关的等式。
   c)有多少种方式从一班 25 个学生中选 6 个学生参加一个委员会?
   d)有多少种方式从一班 25 个学生中选 6 个学生担任委员会中不同的常务委员?
9. a)什么是帕斯卡三角形?
   b)在帕斯卡三角形中的一行是怎样从它的上一行产生的?
10. 什么是恒等式的组合证明?这样的证明与代数证明有什么不同?
11. 解释怎样用组合论证证明帕斯卡恒等式。
12. a)叙述二项式定理。
    b)解释怎样用组合论证证明二项式定理。
    c)求在 $(2x+5y)^{201}$ 的展开式中 $x^{100}y^{101}$ 项的系数。
13. a)解释怎样找出与从 n 个物体允许重复地无序选取 r 个物体的方法数有关的公式。
    b)如果同种类型的物体是不加区分的,那么从 5 种不同类型的物体中选择 12 个物体有多少种方式?
    c)从这 5 种不同类型的物体中选择 12 个物体,如果第一类物体必须至少 3 个,那么有多少种方式?
    d)从这 5 种不同类型的物体中选择 12 个物体,如果第一类物体不多于 4 个,那么有多少种方式?
    e)从这 5 种不同类型的物体中选择 12 个物体,如果第一类物体必须至少 2 个,但是第二类物体不超过 3 个,那么有多少种方式?
14. a)设 n 和 r 是正整数,解释为什么方程 $x_1+x_2+\cdots+x_n=r$ 的解的个数等于 n 元素集合的允许重复的 r 组合数,这里的 x_i 是非负整数,$i=1, 2, 3, \cdots, n$。
    b)方程 $x_1+x_2+x_3+x_4=17$ 有多少个非负整数解?
    c)b)的方程有多少个正整数解?
15. a)n 个物体有 k 种不同的类型,其中类型 1 有 n_1 个无区别的物体,类型 2 有 n_2 个无区别的物体,…,类型 k 有 n_k 个无区别的物体,推导一个与这些物体的排列数有关的公式。
    b)有多少种方式来排序单词 INDISCREETNESS 的字母?
16. 描述一个算法来生成 n 个最小正整数集合的所有排列。
17. a)把 52 张标准的扑克牌发给 6 个人,每人 5 张牌,有多少种方式?
    b)有多少种方式把 n 个有区别的物体分配给 k 个有区别的盒子且使得第 i 个盒子含有 n_i 个物体?
18. 描述一个算法来生成 n 个最小正整数集合的所有的组合?

补充练习

1. 从 10 个不同的项中选 6 项有多少种方式?
   a)若这些项是有序选择的并且不允许重复。　b)若这些项是有序选择的并且允许重复。
   c)若这些项是无序选择的并且不允许重复。　d)若这些项是无序选择的并且允许重复。
2. 从 6 个不同的项中选 10 项有多少种方式?
   a)若这些项是有序选择的并且不允许重复。　b)若这些项是有序选择的并且允许重复。
   c)若这些项是无序选择的并且不允许重复。　d)若这些项是无序选择的并且允许重复。
3. 一个考试包含 100 个真假判断题。如果答案可以空缺,一个学生回答这些考题可能有多少种不同的方式?
4. 有多少个 10 位二进制串以 000 开始或以 111 结束?

5. 字母表$\{a, b, c\}$上有多少个 10 位字符串恰有 3 个 a 或恰有 4 个 b?
6. 一个校园电话系统的内部电话号码由 5 个数字组成，且第一个数字不等于 0。在这个系统中可以分配多少个不同的电话号码?
7. 一个冰激凌屋有 28 种不同味道的冰激凌、8 种不同的果汁和 12 种配料。
   a)如果每种味道的可以不止 1 勺，并且不考虑次序，那么取 3 勺冰激凌放在一个盘中有多少种不同的方式?
   b)如果一个小圣代包含 1 勺冰激凌、1 种果汁和 1 种配料，那么有多少种不同的小圣代?
   c)如果一个大圣代包含 3 勺冰激凌、2 种果汁和 3 种配料。其中每种味道的冰激凌可以不止 1 个并且不考虑次序，每种果汁只能用 1 次且不考虑次序，同时每种配料也只能用 1 次并且不考虑次序。那么有多少种不同的大圣代?
8. 有多少个小于 1000 的正整数
   a)恰有 3 个十进制数字?
   b)有奇数个十进制数字?
   c)至少有 1 个十进制数字等于 9?
   d)没有奇数个十进制数字?
   e)有两个连续的十进制数字等于 5?
   f)是回文(即正读和倒读是一样的)?
9. 当用十进制记法写出从 1 到 1000 的数时，有多少个下面的数字被用到?
   a)0 b)1 c)2 d)9
10. 有黄道十二宫，需要有多少人才能保证其中至少 6 个人在同一宫?
11. 一个幸运卡甜点公司制作 213 种不同的幸运卡。一个学生使用这个公司的幸运卡在餐馆吃饭。如果这个学生没有获得 4 张同样的幸运卡，那么他在这个餐馆吃饭最多可以吃多少次?
12. 为保证至少 2 个人生在一周的同一天和同一个月(可以不在同一年)，那么需要多少人?
13. 证明在 10 个不超过 50 的正整数集合中至少有 2 个不同的 5 元子集有同样的和。
14. 一包棒球卡有 20 张。如果总共有 550 种不同的卡，那么需要买多少包卡才能保证其中的 2 张卡是一样的。
15. a)从一副牌中需要选多少张牌才能保证至少选中 2 张 A?
    b)从一副牌中需要选多少张牌才能保证至少选中 2 张 A 和 2 种点数?
    c)从一副牌中需要选多少张牌才能保证至少有 2 张同样点数的牌?
    d)从一副牌中需要选多少张牌才能保证至少有 2 张不同点数的牌?

*16. 证明在任何 $n+1$ 个不超过 $2n$ 的正整数中必存在 2 个数互素。
*17. 证明在 m 个整数的序列中存在若干个连续的整数其和可被 m 整除。

18. 证明如果放 5 个点在边长为 2 的正方形中，那么其中至少有 2 个点的距离不超过$\sqrt{2}$。
19. 证明一个有理数的十进制展开式一定从某一点出现循环。
20. 一个正 n 边形有多少条对角线? 这里的 n 是大于等于 3 的正整数。
21. 有多少种方式从 20 种多纳圈中选 12 个多纳圈，
    a)如果没有 2 个多纳圈是同种的?
    b)如果所有的多纳圈都是同种的?
    c)如果不加限制?
    d)如果至少有 2 种多纳圈?
    e)如果必须至少有 6 个蓝莓馅的多纳圈?
    f)如果至多有 6 个蓝莓馅的多纳圈?
22. 求 n，如果
    a)$P(n, 2)=110$ b)$P(n, n)=5040$ c)$P(n, 4)=12P(n, 2)$
23. 求 n，如果
    a)$C(n, 2)=45$ b)$C(n, 3)=P(n, 2)$ c)$C(n, 5)=C(n, 2)$
24. 证明如果 n 和 r 是非负整数且 $n\geqslant r$，则
    $P(n+1, r)=P(n, r)(n+1)/(n+1-r)$

*25. 设 S 是 n 元素集合，存在多少个有序对(A, B)使得 A 和 B 是 S 的子集且 $A\subseteq B$? [提示：证明 S 中的

每个元素属于 A，$B-A$ 或 $S-B$。]

26. 通过构造在集合的具有偶数个元素的子集与具有奇数个元素的子集之间的对应，给出关于 4.4 节推论 2 的组合证明。[提示：取定集合的一个元素 a，如下构造对应：如果 a 不在子集中就把它放到子集中；如果 a 在子集中就把它从子集中取出。]
27. 设 n 和 r 是非负整数且 $r<n$。证明
$$C(n,r-1)=C(n+2,r+1)-2C(n+1,r+1)+C(n,r+1)$$
28. 使用数学归纳法证明 $\sum_{j=2}^{n}C(j,2)=C(n+1,3)$，其中 n 是大于 1 的整数。
29. 证明如果 n 是整数，则
$$\sum_{k=0}^{n}3^k\binom{n}{k}=4^n$$
30. 在这个练习中我们将推导一个关于 n 个最小正整数的平方和的公式。我们将用两种方式计数三元组 (i, j, k) 的个数，其中 i，j 和 k 是整数且满足 $0\leqslant i<k$，$0\leqslant j<k$，$1\leqslant k\leqslant n$。

    a) 证明对于给定的 k 存在 k^2 个这样的三元组，因此有 $\sum_{k=1}^{n}k^2$ 个这样的三元组。

    b) 证明具有 $0\leqslant i<j<k$ 的三元组个数和 $0\leqslant j<i<k$ 的三元组个数都等于 $C(n+1, 3)$。

    c) 证明具有 $0\leqslant i=j<k$ 的三元组个数等于 $C(n+1, 2)$。

    d) 把 a)、b) 和 c) 组合起来得出
$$\sum_{k=1}^{n}k^2=2C(n+1,3)+C(n+1,2)$$
$$=n(n+1)(2n+1)/6$$
*31. 设 $n\geqslant 4$，有多少个 n 位二进位串恰好 01 在其中出现两次？
32. 设集合 S 和子集族 A_1，A_2，…，A_n，其中每个子集含有 d 个元素，$d\geqslant 2$。如果可以把两种不同的颜色分配给 S 的元素，每个元素一种颜色，且使得每个子集 A_i 都包含了两种颜色的元素，则称这个子集族是可 2 涂色的。设 $m(d)$ 是最大的正整数，使得对于每个子集族，如果子集数小于 $m(d)$，且每个子集含 d 个元素，就是可 2 涂色的。

    a) 证明具有 $2d-1$ 个元素的集合 S 的所有 d 子集构成的子集族不是可 2 涂色的。

    b) 证明 $m(2)=3$。

    **c) 证明 $m(3)=7$。[提示：证明 {1, 3, 5}，{1, 2, 6}，{1, 4, 7}，{2, 3, 4}，{2, 5, 7}，{3, 6, 7}，{4, 5, 6} 不是可 2 涂色的。然后证明所有具有 3 个元素的 6 个集合的集合族每个都是可 2 涂色的。]
33. 一个教授为一次离散数学考试出了 20 道多选题，每道题可能的答案为 a，b，c 或 d。如果具有答案 a，b，c 和 d 的试题数分别为 8，3，4 和 5，且试题可以用任意的顺序安排，那么可能有多少种不同的答案。
34. 8 个人围圆桌就座有多少种不同的安排？其中如果一种安排通过旋转能从另一种安排得到，那么就认为这两种安排是一样的。
35. 把 24 个学生分给 5 个指导教师有多少种方式？
36. 一蒲式耳包含 20 个无区别的 Delicious 苹果、20 个无区别的 Macintosh 苹果和 20 个无区别的 Granny Smith 苹果，从其中选 12 个苹果，如果每类至少选 3 个有多少种方式？
37. 方程 $x_1+x_2+x_3=17$ 有多少个非负整数解？

    a) 若 $x_1>1$，$x_2>2$，$x_3>3$

    b) 若 $x_1<6$，$x_3>5$

    c) 若 $x_1<4$，$x_2<3$，$x_3>5$
38. 使用单词 PEPPERCORN 的所有字母构成字符串。

    a) 可以构成多少个不同的字符串？

    b) 其中有多少字符串以 P 开始和结束？

c)在多少个字符串中有 3 个连续的 P?

39. 10 元素集合有多少个子集

a)少于 5 个元素? b)多于 7 个元素? c)有奇数个元素?

40. 一个交通逃逸事故的证人告诉警察，肇事汽车的车牌包含 3 个字母后面跟着 3 个数字，以字母 AS 开始且包含数字 1 和 2。有多少不同的车牌符合这个描述?

41. 有多少种方式把 n 个相同的物体放入 m 个不同的容器而使得没有一个容器是空的?

42. 6 个男孩和 8 个女孩坐在一排椅子上，如果没有两个男孩相邻，有多少种方式?

43. 设计一个算法生成一个有穷集的所有允许重复的 r 排列。

44. 设计一个算法生成一个有穷集的所有允许重复的 r 组合。

*45. 证明如果 m，n 为整数，$m\geqslant 3$，$n\geqslant 3$，那么 $R(m, n)\leqslant R(m, n-1)+R(m-1, n)$。

*46. 在一组 6 个人中，每两个人是朋友或者是敌人，证明在这组人中不存在 3 个人两两是朋友，也不存在 4 个人两两是敌人。从而证明了 $R(3, 4)\geqslant 7$。

计算机题目

按下述给定的输入和输出写程序。

1. 给定正整数 n 和不超过 n 的非负整数，找出 n 元素集合的 r 排列数和 r 组合数。
2. 给定正整数 n 和 r，找出 n 元素集合的允许重复的 r 排列数和允许重复的 r 组合数。
3. 给定正整数序列，找出这个序列的最长的递增和递减子序列。

*4. 给定方程 $x_1+x_2+\cdots+x_n=C$，其中 C 是一个常数，x_1，x_2，…，x_n 是非负整数，列出所有的解。

5. 给定正整数 n，按字典顺序列出集合{1，2，…，n}的所有的排列。
6. 给定正整数 n 和不超过 n 的非负整数 r，按字典顺序列出集合{1，2，3，…，n} 的所有的 r 组合。
7. 给定正整数 n 和不超过 n 的非负整数 r，按字典顺序列出集合{1，2，3，…，n} 的所有的 r 排列。
8. 给定正整数 n，列出集合{1，2，…，n}的所有的组合。
9. 给定正整数 n 和 r，列出集合{1，2 ，…，n}的允许重复的所有 r 排列。
10. 给定正整数 n 和 r，列出集合{1，2 ，…，n}的允许重复的所有 r 组合。

计算和研究

使用一个计算程序或你已完成的程序做下面的练习。

1. 当两个队加时赛时，赢的队是 9 分中首先得 5 分、11 分中首先得 6 分、13 分中首先得 7 分和 15 分中首先得 8 分的队。找出加时赛的可能的结果数。
2. 哪些二项式系数是奇数? 你能根据数的特征给出一个猜想吗?
3. 目前还不知道二项式系数 $C(2n, n)$是否一定被一个素数的平方整除，也不知道当 n 增长时在 $C(2n, n)$ 的素数分解中最大的指数是否无界增长。通过对于尽可能多的正整数 n 找$C(2n, n)$的分解式中素数的最小和最大的幂来探索这个问题。
4. 生成 8 元素集合的所有的排列。
5. 生成 9 元素集合的所有的 6 排列。
6. 生成 8 元素集合的所有的组合。
7. 生成 7 元素集合允许重复的所有 5 组合。

写作题目

用本书以外的资料，按下列要求写成短文。

1. 描述狄利克雷和其他的数学家对鸽巢原理的早期应用。
2. 讨论扩充目前电话编码计划的方式以适合对更多电话号码飞速增长的需求。(看看你是否能够找到某些

来自电信产业的建议。)对你要讨论的每个新的编码计划说明怎样找到它所支持的不同电话号码的个数。

3. 本书描述了许多组合恒等式。找一找关于这种恒等式的资料，并且描述除了本书引入之外的其他重要的组合恒等式。给出其中某些恒等式的有代表性的证明，包括组合证明。
4. 描述在统计力学中的质点分布所使用的不同的模型，包括麦克斯韦-玻尔兹曼、玻色-爱因斯坦和费米-狄拉克(Fermi-Dirac)统计，在每种情况下描述模型中使用的计数技术。
5. 定义第一类斯特林数并且描述它们的某些性质以及所满足的恒等式。
6. 定义第二类斯特林数并且描述它们的某些性质以及所满足的恒等式。
7. 定义拉姆齐数，叙述和证明显示它们存在的拉姆齐定理，并且描述目前已知的有关拉姆齐数的结果。
8. 描述生成 n 元素集合所有排列的其他算法，这些算法不是在 4.6 节给出的算法。把这些算法的计算复杂性与书上和 4.6 节练习所描述算法的计算复杂性进行比较。
9. 至少描述一种方法生成一个正整数 n 的所有的剖分。(见 3.4 节练习 47。)



第5章 离散概率

组合学和概率论有着共同的起源。概率论形成于三百年多年以前。当时布莱斯·帕斯卡对某些赌博游戏进行了分析，从中发现了二项式系数的各种性质。后来，法国数学家拉普拉斯[⊖]也研究赌博，他定义了事件的概率，从而带动了基础概率论的大发展。

尽管概率论起源于赌博的研究，但是现在它在各种不同的学科中起着基础的作用。例如，概率论被广泛应用于遗传学的研究，用它可以帮助理解特征的遗传。当然，由于概率论适用于研究人所特别热衷的赌博行为，它仍旧是数学领域里特别流行的一部分。

在计算机科学中，概率论在算法复杂度研究中起着重要的作用。特别地，人们用概率论的思想和技巧确定算法的平均复杂度。概率算法可以用于解决许多不容易或实际上不可能用确定性算法求解的问题。确定性算法在给定同样的输入条件以后，总是遵循着同样的步骤，但在概率算法中不是这样，算法做一次或多次随机选择，这可能导致不同的输出结果。在组合学中，概率论甚至可以用于证明具有特定性质的个体的存在性。由保罗·埃德斯和阿尔弗雷德·任伊引入组合学的概率方法，通过证明存在具有某种性质个体的概率是正数来证明这种个体的存在性。

5.1 离散概率引论

5.1.1 引言

概率论起源于17世纪的赌博游戏，人们根据反复掷一对骰子的结果来赌输赢。法国数学家布莱斯·帕斯卡计算了几种常见赌法中赢的可能性。到了18世纪，法国数学家拉普拉斯也研究赌博，并且把事件的概率定义为成功的结果数除以可能的结果数所得的商。例如，一个骰子掷出奇数点的概率就是成功结果的个数(即出现奇数点的个数)除以可能结果的个数(即骰子可能出现的不同方式数)。有6种可能的结果，即1，2，3，4，5和6，其中恰好3种是成功的结果，即1，3和5。因此，骰子掷出奇数点的概率是3/6=1/2。(注意这里假定所有结果的可能性是相等的，或者换句话说，骰子是均匀的。)

这一节我们的讨论将局限于具有可能性相等的有限多个结果的试验。这样我们可以使用拉普拉斯关于事件概率的定义。我们将在5.2节研究具有有限多个结果但结果的可能性不一定相等的试验，我们将引入概率论中的一些关键概念，包含条件概率、事件的独立性、随机变量。在5.3节将引入一个随机变量的期望和方差的概念。



⊖ **皮埃尔-西蒙·拉普拉斯(Pierre-Simon Laplace，1749—1827)** 拉普拉斯出身于诺曼底的一个下层家族，童年时在一所教会学校受教育。16岁他进入凯恩大学学习神学，但是不久他意识到自己真正感兴趣的是数学。毕业后，他在凯恩大学担任临时教授。在1769年他成为巴黎陆军学校的数学教授。

拉普拉斯由于对天体力学、天体运动研究所做出的贡献而闻名于世。他的《天体力学》被认为是19世纪初期最伟大的科学著作之一。拉普拉斯是概率论的奠基人之一，还对数理统计学做出了许多贡献。他把在这个领域的工作写成《概率论的理论分析》一书，书中定义一个事件的概率为试验所希望的结果数与总结果数之比。

拉普拉斯还由于他的政治灵活性而著称。他先后忠实于法兰西共和国、拿破仑和路易十八国王。这使他在法国革命前、革命期间和革命后都成果卓著。

5.1.2 有限概率

我们把从一组可能的结果中得出一个结果的过程称为试验。试验的样本空间是可能结果的集合。一个事件是样本空间的子集。现在叙述拉普拉斯关于具有有限多个可能结果的事件的概率定义。

定义 1 事件 E 是结果具有相等可能性的有限样本空间 S 的子集，则事件 E 的概率是 $p(E)=|E|/|S|$。

例 1～例 8 说明怎样找出事件的概率。

Extra Examples

例 1 缸里有 4 个蓝球和 5 个红球。从缸里取出一个蓝球的概率是多少？

解 为计算这个概率，首先考虑存在 9 个可能的结果，这些可能的结果中有 4 个得到蓝球。因此，选一个蓝球的概率是 4/9。■

例 2 掷两个骰子使得其点数之和等于 7 的概率是多少？

解 当掷两个骰子时总共有 36 种可能的结果(这是由乘积法则得到的，因为每个骰子有 6 个可能的结果，当掷两个骰子时总共有 $6^2=36$ 种结果)。存在有 6 种成功的结果，即(1，6)，(2，5)，(3，4)，(4，3)，(5，2)和(6，1)，这里两个骰子的点数用一个有序对来表示。因此，掷两个均匀的骰子时，点数和为 7 出现的概率是 6/36＝1/6。■

Links

目前彩票非常流行。我们可以轻松地算出赢各种不同类型彩票的机会。

例 3 在一种彩票里，人们挑 4 个数字，如果数字与一个随机机械过程选出的 4 个数字吻合且次序相同，他们就中了大奖。如果只有 3 个数字匹配，他们就中了比较小的奖。那么，赢大奖的概率是多少？赢小奖的概率是多少？

解 选择的 4 个数字都正确的方法只有一种。而由乘积法则，任选 4 个数字共有 $10^4=10\,000$ 种方式。因此，赢大奖的概率是 1/10 000＝0.0001。

4 个数字中恰好选对了 3 个数字的能够赢小奖。为了使得 3 个数字正确，而不是 4 个数字全对，必须恰好 1 个数字出错。可以先求选 4 个数字且除了第 i 个数字之外都与挑出的数字匹配的方式数，这里的 $i=1，2，3，4$，然后对它们求和。根据求和法则，就能得到恰好选对 3 个数字的方式数。先求第 1 个数字不匹配的选法数，观察到对第 1 个数字有 9 种可能的选择(除了一个正确的数字外)，而其他的每个数字只有一种选择，即对应位置的正确数字。因此，第 1 个数字出错而后 3 个数字正确的选法有 9 种。类似地，有 9 种方式选出 4 个数字而只有第 2 个数字出错，又有 9 种方式只有第 3 个数字出错，以及 9 种方式只有第 4 个数字出错。从而总共有 36 种方式选择 4 个数字，并恰好其中 3 个是正确的。于是，赢得小奖的概率是 36/10 000＝9/2500＝0.0036。■

例 4 现在有许多彩票要求从 1 到正整数 n 中选出 6 个数的数组，选对的人得到特别大奖，这里的 n 通常在 30 到 60 之间。一个人从 40 个数中选对 6 个数的概率是多少？

解 只有一个赢奖的组合，从 40 个数中选 6 个数的总方法数是

$$C(40,6)=\frac{40!}{34!6!}=3\,838\,380$$

因此，选出一个赢奖组合的概率是 1/3 838 380≈0.000 000 26(符号≈表示近似等于)。■

Links

我们可以使用现有的手段找到扑克游戏中某些牌的概率。一副牌有 52 张，分成 13 个不同的类，每类 4 张。这些类是 2，3，4，5，6，7，8，9，10，J，Q，K 和 A。还有 4 套花色，黑桃、梅花、红心和方块，每套包含 13 张，并且在一种花色中每类只有 1 张牌。

例 5 如果 5 张为一手牌，从 52 张牌中可选出多少手不同的牌？

解 存在 $C(52, 5)=2\,598\,960$ 种不同的牌。 ■

例 6 一手扑克牌有 5 张，如果要求其中的 4 张牌是同一类的，找出这种牌的概率。

解 由乘积法则，5 张牌中 4 张牌是同类的数目，就是任选一类的方式数、从该类的 4 张牌中选取 4 张牌的方式数与选第 5 张牌的方式数之积，即

$$C(13,1)C(4,4)C(48,1)$$

由于总共存在 $C(52, 5)$ 手牌，因此一手牌中包含 4 张同类牌的概率是

$$\frac{C(13,1)C(4,4)C(48,1)}{C(52,5)}=\frac{13\cdot 1\cdot 48}{2\,598\,960}\approx 0.000\,24$$ ■

例 7 一手牌打出满堂红，即 3 张在同一类且其余 2 张在另一类的概率是多少？

解 由乘积法则，打出满堂红的方式数也就是有序地选取两个类的方式数，即第一类的 4 张牌选 3 张的方式数和第二类的 4 张牌选 2 张的方式数之积(注意两类的次序是有关的，例如 3 个 Q 和 2 个 A，与 3 个 A 和 2 个 Q 是不同的。)可以看出打出满堂红的方式数是

$$P(13,2)C(4,3)C(4,2)=13\cdot 12\cdot 4\cdot 6=3744$$

因为存在 2 598 960 手牌，所以出现满堂红的概率是

$$\frac{3744}{2\,598\,960}\approx 0.0014$$ ■

例 8 箱子里有 50 个球，依次标号为 1，2，…，50，依次取出号码为 11，4，17，39，23 的球的概率是多少？如果(a)在选下一个球之前已经选出的球不再放回到箱子里；(b)在选下一个球之前已经选出的球要放回到箱子里。

解 (a)根据乘积法则，存在 $50\cdot 49\cdot 48\cdot 47\cdot 46=254\,251\,200$ 种方法选球。因为每当一个球拿走，就少一个被选的球。因此 11，4，17，39，23 号球被依次取出的概率是 1/254 251 200。这是一个*无放回抽样*的实例。

(b)根据乘积法则，存在 $50^5=312\,500\,000$ 种方法选球，因为每次拿走一个球，还存在 50 种可选的球。因此 11，4，17，39，23 号球被依次取走的概率是 1/312 500 000。这是一个*有放回抽样*的实例。 ■

5.1.3 事件组合的概率

我们可以使用计数方法得到从其他事件导出的事件的概率。

定理 1 设 E 是样本空间 S 的一个事件。事件 $\bar{E}$(事件 E 的补事件)的概率是

$$p(\bar{E})=1-p(E)$$

证 为求出事件 $\bar{E}$ 的概率，我们注意到 $|\bar{E}|=|S|-|E|$。因此，

$$p(\bar{E})=\frac{|S|-|E|}{|S|}=1-\frac{|E|}{|S|}=1-p(E)$$ □

当直接的方法不适用时，可以采取其他办法来找事件的概率。不用直接求这个事件的概率，但可以确定它的补事件的概率。这往往更容易做到，正如下面的例子所示。

例 9 随机生成一个 10 位数的二进位序列，其中至少 1 位是 0 的概率是多少？

解 设 E 是 10 位中至少一位是 0 的事件。那么 $\bar{E}$ 是所有的位都是 1 的事件。因为样本空间是所有 10 位二进位串的集合，从而得到

$$p(E)=1-p(\bar{E})=1-\frac{|\bar{E}|}{|S|}=1-\frac{1}{2^{10}}=1-\frac{1}{1024}=\frac{1023}{1024}$$

所以，包含至少一位 0 的二进位串的概率是 1023/1024。不用定理 1 而直接求这个概率是相当困难的。 ■

我们也可以求出两个事件的并集的概率。

定理 2 设 E_1 和 E_2 是样本空间的事件，那么

$$p(E_1 \cup E_2) = p(E_1) + p(E_2) - p(E_1 \cap E_2)$$

证 使用 1.7 节给出的关于两个集合的并集的元素数公式得到

$$|E_1 \cup E_2| = |E_1| + |E_2| - |E_1 \cap E_2|$$

因此，

$$\begin{aligned} p(E_1 \cup E_2) &= \frac{|E_1 \cup E_2|}{|S|} \\ &= \frac{|E_1| + |E_2| - |E_1 \cap E_2|}{|S|} \\ &= \frac{|E_1|}{|S|} + \frac{|E_2|}{|S|} - \frac{|E_1 \cap E_2|}{|S|} \\ &= p(E_1) + p(E_2) - p(E_1 \cap E_2) \end{aligned}$$

□

例 10 从不超过 100 的正整数中随机选出一个正整数，它能被 2 或 5 整除的概率是多少？

解 设 E_1 是选出一个能被 2 整除的数的事件，E_2 是选出一个能被 5 整除的数的事件。那么 $E_1 \cup E_2$ 是能被 2 或 5 整除的事件，$E_1 \cap E_2$ 是能被 2 和 5 同时整除的事件，即能被 10 整除的事件。由于 $|E_1| = 50$，$|E_2| = 20$，且 $|E_1 \cap E_2| = 10$，从而得到

$$\begin{aligned} p(E_1 \cup E_2) &= p(E_1) + p(E_2) - p(E_1 \cap E_2) \\ &= \frac{50}{100} + \frac{20}{100} - \frac{10}{100} = \frac{3}{5} \end{aligned}$$

■

5.1.4 概率的推理

一个常见的问题是确定两个事件中的哪一个更有可能发生，分析这些事件的概率可能比较复杂。下面的例子描述了一个这样的问题，它讨论了一个来自电视游戏节目《让我们成交》的著名问题。

例 11 蒙地厅大厦的 3 门难题 假定你参与游戏，有机会赢一个大奖。参与者从 3 扇门中选一扇打开，大奖只在其中一扇门的后面。节目主持人知道每扇门后面是什么，一旦你选中了某扇门，不管是否选择了中奖的门，他都会打开另外一扇没有奖的门(如果两扇门后都没有奖，就随便打开一扇)。然后他问你是否愿意换另外一扇门。你应该用什么策略？你应该换一扇门，还是坚持原来的选择，或者这无关紧要？

解 在主持人开门之前，你选对了门的概率是 1/3，因为这 3 扇门中奖的可能性相等。等主持人打开另外一扇门之后，你所选的门正确的概率不变，因为他总是打开后面没有大奖的门。

你选错门的概率就是大奖在你没有选的两扇门中某一扇后面的概率，因此，你选错了门的概率是 2/3。如果你选错了，这时主持人打开一扇门向你显示大奖不在它的后面，大奖一定在另一扇门之后。若你原来的选择是错的，这时改变主意，那么你总能赢。因此，通过改变选择，你赢的概率是 2/3。也就是说，当主持人给你这样做的机会时，你总应该选择改变，这使得你赢的概率增加了一倍。 ■

练习

1. 从一副牌中选出 1 张 A 的概率是多少？
2. 掷骰子时出现 6 点的概率是多少？
3. 从前 100 个正整数中随机选出 1 个奇数的概率是多少？

4. 从一年(366天)中随机选出1天在4月的概率是多少?
5. 当掷2个骰子时，其点数之和是偶数的概率是多少?
6. 从一副牌中选1张牌是张A或是张红心的概率是多少?
7. 掷6次硬币，全部头像向上的概率是多少?
8. 一手扑克牌有5张，其中包含红心A的概率是多少?
9. 一手扑克牌有5张，其中不包含红心Q的概率是多少?
10. 一手扑克牌有5张，其中包含方块2和黑桃3的概率是多少?
11. 一手扑克牌有5张，其中包含方块2、黑桃3、红心6、梅花10和红心K的概率是多少?
12. 一手扑克牌有5张，其中恰好包含1张A的概率是多少?
13. 一手扑克牌有5张，其中至少包含1张A的概率是多少?
14. 一手扑克牌有5张，其中包含5类不同的牌的概率是多少?
15. 一手扑克牌有5张，其中包含2个对子(两张牌花色不同但类相同)的概率是多少?
16. 一手扑克牌有5张，其中包含一手同花，即5张牌的花色相同的概率是多少?
17. 一手扑克牌有5张，其中包含一个顺子，即5张牌的类是连续的概率是多少?[注意：A—2—3—4—5和10—J—Q—K—A都可以看成是顺子。]
18. 一手扑克牌有5张，其中包含一个同花顺子，即5张牌的类连续且是同一花色的概率是多少?
*19. 一手扑克牌有5张，其中包含5张不同类的牌且不包含一个同花或一个顺子的概率是多少?
20. 一手扑克牌有5张，其中包含同一花色的10，J，Q，K和A的概率是多少?
21. 一个骰子掷6次都不出现偶数点的概率是多少?
22. 随机选取一个不超过100的正整数，能够被3整除的概率是多少?
23. 随机选取一个不超过100的正整数，能够被5或7整除的概率是多少?
24. 求从不超过下述整数的正整数中选中6个整数来赢彩票的概率，这里不考虑选择整数的顺序。
 a)30　b)36　c)42　d)48
25. 求从不超过下述整数的正整数中选中6个整数来赢彩票的概率，这里不考虑选择整数的顺序。
 a)50　b)52　c)56　d)60
26. 求从不超过下述整数的正整数中选6个整数都不中的概率，这里不考虑选择整数的顺序。
 a)40　b)48　c)56　d)64
27. 求从不超过下述整数的正整数中选6个整数，并且恰好选中1个的概率，这里不考虑选择整数的顺序。
 a)40　b)48　c)56　d)64
28. 美国宾夕法尼亚超级彩票的玩法是，买彩票的人要从前80个正整数中选出7个数。如果这7个数是在由宾夕法尼亚彩票委员会选出的11个数之6中就能赢大奖，那么一个人赢大奖的概率是多少?
29. 在一种超级彩票中，如果买彩票的人选中的8个数正是计算机从不超过100的正整数中选出的数就能中彩。请问中彩的概率是多少?
30. 由计算机从1到40之间(包括1和40在内)选出6个数，如果某人选中了其中的5个(但不是6个)数就能获奖，那么获奖的概率是多少?
31. 假设100个人进入决赛并且随机选择不同的人作为一等奖、二等奖和三等奖的获奖者。如果米切尔是进入决赛的人之一，她中奖的概率是多大?
32. 假设100个人进入决赛并且随机选择不同的人作为一等奖、二等奖和三等奖的获奖者。如果库玛、加奈斯、彼得罗是进入决赛的人，他们每个人都赢得一个奖项的概率是多少?
33. 在一次绘画比赛中，200个人进入决赛，在下述条件下，艾比、巴里、西尔维亚分别赢得一等奖、二等奖、三等奖的概率是多少?
 a)如果每个人至多得一个奖。　b)如果允许一个人得多个奖。
34. 在一次绘画比赛中，50个人进入决赛，在下述条件下，勃、考林、杰夫、罗海尼分别赢得一等奖、二等奖、三等奖和四等奖的概率是多少?
 a)如果每个人至多得一个奖。　b)如果允许一个人得多个奖。

35. 在轮盘赌中，旋转一个有38个数的轮盘，其中18个数是红的，18个数是黑的，另外两个既不红也不黑的数是0和00。当轮盘转动时，它到达任何特定数字的概率是1/38。
    a)轮盘落到1个红数的概率是多少？
    b)轮盘两次落到某列上的同一个黑数的概率是多少？
    c)轮盘落到0或00的概率是多少？
    d)轮盘旋转5次，5次都不落到0或00的概率是多少？
    e)某次转动轮盘，落到1和6之间(包含1和6在内)的某个数字，但下次转动轮盘却不落到这些数字之间的概率是多少？
36. 掷2个骰子总点数为8或掷3个骰子总点数为8，哪种可能性更大？
37. 掷2个骰子总点数为9或掷3个骰子总点数为9，哪种可能性更大？
38. 设 E_1 和 E_2 是两个事件，如果 $p(E_1 \cap E_2) = p(E_1)p(E_2)$，就称 E_1 和 E_2 是*独立的*。如果把一枚硬币被抛掷3次时所有可能的结果构成一个集合，把这个集合的子集看作事件，确定下面的每一对事件是否是独立的。
    a) E_1：第一次硬币头像向下；E_2：第二次硬币头像向上。
    b) E_1：第一次硬币头像向下；E_2：在连续3次中有2次但不是3次头像向上。
    c) E_1：第二次硬币头像向下；E_2：在连续3次中有2次但不是3次头像向上。
    (我们将在5.2节更深入地研究事件的独立性。)
39. 解释下面的陈述错在什么地方。在蒙地厅大厦三门难题里，因为剩下了两个门，你选的第一个门后面是大奖的概率与另一个没打开的门后面是大奖的概率都是1/2。
40. 假定在蒙地厅大厦难题中不是三个门而是四个门。当知道每个门后面是什么的主持人打开一个后面并没有奖品的门并且给你机会改变选择时，你不改变选择并且赢了大奖的概率是多少？在还剩下两个门没有打开时，你改变原来的选择猜中两个门其中一个后面有奖的概率是多少？
41. 这个问题由薛瓦利埃·德梅雷[⊖]提出，并由布莱斯·帕斯卡和皮埃尔·德·费马解决。
    a)求一个骰子掷4次时掷出一个6点的概率。
    b)求一对骰子掷24次时掷出两个6点的概率。这个问题是薛瓦利埃·德梅雷问帕斯卡的，他问这个概率是否大于1/2。请解答这个问题。
    c)一个骰子掷4次时掷出一个6点或一对骰子掷24次时掷出两个6点，哪种情况更可能发生？

5.2 概率论

5.2.1 引言

在5.1节我们引入了事件的概率的概念。(回忆一下，一个事件是一次试验的可能结果的子集。)我们按照拉普拉斯的记号定义一个事件 E 的概率

$$p(E) = \frac{|E|}{|S|}$$

即 E 中的结果个数除以结果总数。这个定义假定所有结果的可能性都是相等的。但是许多试验结果的可能性并不相等。例如，一个硬币很可能是不均匀的，因而出现头像向上的次数常常是向下次数的两倍。类似地，一个线性搜索的输入是一个元素和一个表，这个元素在表里或不在表里的可能性依赖于输入的产生过程。在这种情况下，怎样建立关于事件可能性的模型呢？这一节将要说明当结果的可能性不相等时，为研究试验概率应该怎样定义结果的概率。

⊖ 薛瓦利埃·德梅雷(Chevalier de Méré)是法国贵族，他以赌博闻名于世，也是一个纨袴子弟。他善于打赌，胜率都略大于1/2(例如在一个骰子掷4次时至少出现一个6点)。他给帕斯卡写信问到关于一对骰子掷24次至少出现两个6点的概率，这带动了概率论的发展。据说，帕斯卡写信给费马谈到薛瓦利埃，说过诸如"他是一个好人，但是，唉，他不是数学家。"这样的话。

假定一个均匀的硬币被掷 4 次，第一次它的头像向上。给定了这个信息，头像 3 次向上的概率是多少？为了回答这个问题或者类似的问题，我们将引入条件概率的概念。已知第一次头像向上能改变 3 次头像向上的概率吗？如果不是，这两个事件就叫做独立的，本节的后面将要学到这个概念。

许多问题谈到一个与试验结果有关的特定数值。例如，当我们掷 100 次硬币时，恰好出现 40 次头像的概率是多少？我们应该预期出现多少次头像？在这一节我们将要学习随机变量，它是把数值与试验结果联系起来的函数。

5.2.2 概率指派

设 S 是某个具有有穷个或可数个结果的试验的样本空间。我们赋给每个结果 s 一个概率 $p(s)$，使得满足以下两个条件：

(i) $0 \leqslant p(s) \leqslant 1$　对每个 $s \in S$

和

(ii) $\sum_{s \in S} p(s) = 1$

条件(i)说明每个结果的概率是一个不超过 1 的非负实数。条件(ii)说明所有可能结果的概率之和应该是 1，即当我们做这个试验时这些结果之一一定出现。这是拉普拉斯定义的一般化。在拉普拉斯定义中，n 个结果中的每一个，其概率都是 $1/n$。的确，当使用拉普拉斯关于等可能结果概率的定义且 S 有限时，条件(i)和(ii)是满足的(见练习 4)

注意到当存在 n 个可能的结果 x_1，x_2，…，x_n 时，这两个要满足的条件是

(i) $0 \leqslant p(x_i) \leqslant 1$　对 $i=1, 2, \cdots, n$

和

(ii) $\sum_{i=1}^{n} p(x_i) = 1$

样本空间 S 的所有事件的集合上的函数 p 称作概率分布。

为了建立试验的模型，赋给结果 s 的概率 $p(s)$ 应该等于 s 出现次数除以试验进行的次数。当这个数无限增加时，就取极限。(我们将假定讨论的所有试验有着平均可预料的结果，以使得这个极限存在。我们也假定一个试验的结果成功与否与前面的结果无关。)

注意　我们将只讨论结果集合离散时的事件概率，而一个试验的结果可能是任何实数这样的情况不讨论。此时，对于事件概率的研究通常要求微积分。

我们可以建立试验的模型，在这种试验中结果具有等可能性，或者不等但可以选择一个适当的函数 $p(s)$ 来表示，正如例 1 所示。

例 1　当一个均匀的硬币被掷时，结果 H(头像向上)和结果 T(头像向下)应该赋予什么概率？当硬币不均匀而使得出现头像向上的次数常常是向下的两倍，对这些事件又应该赋予什么概率？

解　对于均匀的硬币，当硬币被掷时头像向上的概率等于头像向下的概率，这两个事件是等可能的。因此，我们对这两个可能结果中的任何一个都赋予 1/2 的概率，即 $p(H)=p(T)=1/2$。

对于不均匀的硬币有

$$p(H) = 2p(T)$$

由于

$$p(H)+p(T)=1$$

从而得出

$$2p(T)+p(T)=3p(T)=1$$

最终有 $p(T)=1/3$ 和 $p(H)=2/3$。 ■

定义 1 假设 S 是 n 元素集合。均匀分布赋给 S 中的每个元素的概率是 $1/n$。

注意均匀分布时，事件的概率与针对拉普拉斯初始概率定义事件的概率是相同的。从一个具有均匀分布的样本空间选一个元素的试验称为随机选取一个元素。

现在我们把事件的概率定义成在这个事件中结果的概率之和。

定义 2 事件 E 的概率是在 E 中结果的概率之和，即

$$p(E)=\sum_{s\in E}p(s)$$

注意当事件 E 中有 n 个结果时，即如果 $E=\{a_1,a_2,\cdots,a_n\}$，则 $p(E)=\sum_{i=1}^{n}p(a_i)$。

例 2 假定一个骰子是不均匀的(或经装填的)，使得 3 这一面出现的次数是其他面的两倍，但其他五个面出现是等可能的。当我们掷这个骰子时，出现奇数点的概率是多少?

解 我们想要找到事件 $E=\{1, 3, 5\}$ 的概率。由本节末的练习 2，我们有

$$p(1)=p(2)=p(4)=p(5)=p(6)=1/7;\ p(3)=2/7$$

从而得出

$$p(E)=p(1)+p(3)+p(5)=1/7+2/7+1/7=4/7$$ ■

当事件是等可能的并且存在有限多个可能的结果时，在这一节给出的事件概率的定义(定义 2)与拉普拉斯的定义(5.1 节定义 1)一致。为此，假定存在 n 个等可能的结果；由于这些概率之和是 1，因此每个可能结果的概率是 $1/n$。假定事件 E 包含 m 个结果，根据定义 2，

$$p(E)=\sum_{i=1}^{m}\frac{1}{n}=\frac{m}{n}$$

由于 $|E|=m$ 和 $|S|=n$，从而有

$$p(E)=\frac{m}{n}=\frac{|E|}{|S|}$$

这是事件 E 的拉普拉斯的概率定义。

5.2.3 事件的组合

当使用定义 2 来定义事件概率时，在 5.1 节中事件组合的概率公式继续保持。例如，5.1 节定理 1 断言

$$p(\overline{E})=1-p(E)$$

其中 $\overline{E}$ 是事件 E 的补事件。当用定义 2 时这个等式也成立。为此只需注意到 n 个可能结果的概率之和是 1，且每个结果或在 E 或在 $\overline{E}$ 中，但不能同时在两者之中。因而

$$\sum_{s\in S}p(s)=1=p(E)+p(\overline{E})$$

所以，$p(\overline{E})=1-p(E)$。

根据拉普拉斯的定义，由 5.1 节定理 2，我们有

$$p(E_1\cup E_2)=p(E_1)+p(E_2)-p(E_1\cap E_2)$$

其中 E_1 和 E_2 是样本空间 S 的事件。当我们按照这一节的做法定义事件的概率时，等式也成立。为此，注意到 $p(E_1\cup E_2)$ 是在 $E_1\cup E_2$ 中结果的概率之和。当结果 x 只属于 E_1 和 E_2 中的一个集

合但不同时属于两个集合时，$p(x)$恰好只出现在 $p(E_1)$或 $p(E_2)$的一个和里。当结果 x 同时在 E_1 和 E_2 中时，$p(x)$出现在 $p(E_1)$的和里，$p(E_2)$的和里，也出现在 $p(E_1 \cap E_2)$的和里。因此它在右边出现了 $1+1-1=1$ 次。所以，左边与右边相等。

可以用定理 1 找到两两不交的事件的并的概率。

定理 1 如果 E_1，E_2，…是样本空间 S 中两两不交的事件的序列，那么

$$p\left(\bigcup_i E_i\right) = \sum_i p(E_i)$$

（注意当序列 E_1，E_2，…由有穷个或可数无穷个两两不交的事件组成时，定理仍旧适用。）

定理的证明留给读者完成(见练习 36 和 37)。

5.2.4 条件概率

假定我们掷 3 次硬币，并且所有的 8 种结果都是等可能的。此外，假定我们知道第一次掷硬币头像向下的事件 F 已经出现了。在给定这一信息后，事件 E，即头像向下出现奇数次的概率是什么？因为第一次掷硬币的头像向下，只有 4 种可能的结果：TTT，TTH，THT 和 THH，其中 H 和 T 分别表示头像向上和向下。头像向下出现奇数次的情况只有 TTT 和 THH。由于 8 个结果的概率相等，在给定 F 出现的条件下，4 种可能的结果的每一个也应该有相等的概率 1/4。这就告诉我们，在给定 F 出现的条件下，E 的概率应为 $2/4=1/2$。这个概率叫做给定 F 的条件下 E 的条件概率。

一般说来，为了找出给定 F 的条件下 E 的条件概率，我们用 F 作为样本空间。作为要出现的 E 的一个结果，这个结果也必须属于 $E \cap F$。由此，我们得到下述定义。

定义 3 设 E 和 F 是具有 $p(F)>0$ 的事件，给定 F 的条件下 E 的条件概率记作 $p(E \mid F)$，定义为

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)}$$

Extra Examples

例 3 随机生成 4 位的二进制串以使得 16 个位串都是等可能的，那么在给定串的第一位是 0 的条件下，串中至少含有两个连续的 0 的概率是多少？（假定 0 位和 1 位是等可能的。）

解 设 E 事件是 4 位二进制串至少含有 2 个连续的 0，F 事件是 4 位二进制串的第一位是 0；那么在给定第一位是 0 的条件下，4 位二进制串包含至少 2 个连续 0 的概率是

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)}$$

由于 $E \cap F=\{0000, 0001, 0010, 0011, 0100\}$，故 $p(E \cap F)=5/16$。因为以 0 开始的 4 位二进制串有 8 个，因而 $p(F)=8/16=1/2$。所以，

$$p(E \mid F) = \frac{5/16}{1/2} = \frac{5}{8}$$ ■

例 4 在至少已有 1 个男孩的条件下，一个家庭中的两个孩子均是男孩的条件概率是多少？假定 BB，BG，GB 和 GG 是等可能的，其中 B 代表男孩，G 代表女孩。

解 设 E 是两个孩子均是男孩的概率，F 是两个孩子中至少有一个是男孩的概率。因而 $E=\{BB\}$，$F=\{BB, BG, GB\}$，并且 $E \cap F=\{BB\}$。由于 4 种可能性是等可能的，故 $p(F)=3/4$ 且 $p(E \cap F)=1/4$。从而可以断言

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)} = \frac{1/4}{3/4} = \frac{1}{3}$$ ■

5.2.5 独立性

假设一个硬币被掷了3次，正如我们关于条件概率讨论的引言中所描述的。第一次掷出的头像向下(事件F)是否改变了头像向下(事件E)次数为奇数的概率？换句话说，$p(E|F)=p(E)$？由于$p(E|F)=1/2$和$p(E)=1/2$，这个等式对事件E和F是有效的。因为这个等式成立，我们说E和F是**独立的事件**。

由于$p(E|F)=p(E\cap F)/p(F)$，是否有$p(E|F)=p(E)$等价于是否有$p(E\cap F)=p(E)p(F)$，从而得到定义4。

定义4 事件E和F是独立的当且仅当$p(E\cap F)=p(E)p(F)$。

例5 假设E是随机产生以一个1开始的4位二进制串的事件，F是随机产生包含偶数个0的二进位串的事件。如果16个4位二进位串是等可能的，E和F是独立的吗？

解 以1开始的4位二进位串有8个：1000，1001，1010，1011，1100，1101，1110和1111。包含偶数个0的4位二进位串也有8个：0000，0011，0101，0110，1001，1010，1100，1111。因为4位二进位串有16个，故

$$p(E)=p(F)=8/16=1/2$$

由于$E\cap F=\{1111, 1100, 1010, 1001\}$，因此

$$p(E\cap F)=4/16=1/4$$

因为

$$p(E\cap F)=1/4=(1/2)(1/2)=p(E)p(F)$$

我们断定E和F是独立的。 ■

概率在遗传学上也有许多应用，如例6和例7所示。

例6 和例4类似，假定一个家庭两个孩子有4种等可能的情况。事件E是有两个孩子的家庭有两个男孩，事件F是有两个孩子的家庭至少有一个男孩，E和F是否独立？

解 因为$E=\{BB\}$，我们有$p(E)=1/4$。在例4我们证明了$p(F)=3/4$和$p(E\cap F)=1/4$。由于$p(E\cap F)=1/4\neq 3/16=(1/4)(3/4)=p(E)p(F)$，事件$E$和$F$不是独立的。 ■

例7 事件E是某个有三个孩子的家庭有男孩也有女孩，F是有三个孩子的家庭至多有一个男孩。假定一个家庭可能有三个孩子的8种方式是等可能的，E和F是否独立？

解 一个家庭可能有三个孩子的8种方式是BBB，BBG，BGB，BGG，GBB，GBG，GGB，GGG，每一种的概率都是1/8。因为$E=\{BBG, BGB, BGG, GBB, GBG, GGB\}$，$F=\{BGG, GBG, GGB, GGG\}$，并且$E\cap F=\{BGG, GBG, GGB\}$，从而$p(E)=6/8=3/4$，$p(F)=4/8=1/2$，且$p(E\cap F)=3/8$。由于

$$p(E\cap F)=\frac{3}{8}=\frac{3}{4}\cdot\frac{1}{2}=p(E)p(F)$$

可以断言E和F是独立的。(这个结论似乎是令人惊奇的。的确，如果我们改变孩子的数目，结论可能不再成立。见节末的练习27。) ■

5.2.6 伯努利试验与二项分布



假设一个试验只有两种可能的结果。例如，当随机产生一个二进位时，可能的结果就是0和1。当一个硬币被掷时，可能的结果就是头像向上和头像向下。每次实行一项

具有两种可能结果的试验就叫做一次伯努利试验。它是以詹姆斯·伯努利⊖命名的，他对概率论作出了重要的贡献。一般来说，一次伯努利试验的一个可能的结果叫做**成功**或**失败**。如果 p 是一次成功的概率，q 是一次失败的概率，那么 $p+q=1$。

当一个试验由 n 次独立的伯努利试验组成时，许多问题可以通过确定 k 次成功的概率来解决。考虑下面的例子。

例 8 一枚硬币是不均匀的以至于出现头像的概率是 2/3。假定每次掷硬币是独立的，当掷 7 次硬币时恰好 4 次出现头像的概率是多少？

解 当一枚硬币被掷 7 次时存在 $2^7=128$ 种可能的结果，7 次中有 4 次出现头像的方式数是 $C(7,4)$。因为 7 次掷币是独立的，每一个这样的结果都对应概率 $(2/3)^4(1/3)^3$。因此，恰好 4 次出现头像的概率是：

$$C(7,4)(2/3)^4(1/3)^3=\frac{35\cdot 16}{3^7}=\frac{560}{2187} \qquad \blacksquare$$

参照在例 8 中用过的同样的推理，我们可以找出在 n 次独立的伯努利试验中有 k 次成功的概率。

定理 2 *在 n 次独立的伯努利试验中有 k 次成功的概率* 在成功概率为 p、失败概率为 $q=1-p$ 的 n 次独立的伯努利试验中，有 k 次成功的概率是：

$$C(n,k)p^kq^{n-k}$$

证 当执行 n 次伯努利试验时，结果是 n 元组 $(t_1, t_2, \cdots, t_n)$，其中 $t_i=S$(成功)或 $t_i=F$(失败，$i=1, 2, \cdots, n$。由于 n 次试验是独立的，由 k 次成功和 $n-k$ 次失败(以任何顺序)组成的每个 n 次试验结果的概率是 p^kq^{n-k}。因为由 S 和 F 构成的包含 k 个 S 的 n 元组有 $C(n,k)$ 个，k 次成功的概率是：

$$C(n,k)p^kq^{n-k} \qquad \square$$

我们将成功概率为 p、失败概率为 $q=1-p$ 的 n 次独立的伯努利试验中有 k 次成功的概率记作 $b(k;n,p)$。作为 k 的函数，我们把这个函数称为**二项分布**。定理 2 告诉我们，$b(k;n,p)=C(n,k)p^kq^{n-k}$。

Extra Examples

例 9 当产生 10 位二进制串时，若每一位为 0 的概率是 0.9，为 1 的概率是 0.1，且每一位的产生是独立的，那么恰好产生 8 位 0 的概率是多少？

解 由定理 2，恰好产生 8 位 0 的概率是

$$b(8;10,0.9)=C(10,8)(0.9)^8(0.1)^2=0.193\,710\,244\,5 \qquad \blacksquare$$

注意当执行 n 次独立的伯努利试验时，对于 $k=0, 1, 2, \cdots, n$，存在 k 次成功的概率之和等于

$$\sum_{k=0}^{n}C(n,k)p^kq^{n-k}=(p+q)^n=1$$

Links ⊖ **詹姆斯·伯努利(James Bernoulli, 1654—1705)** 伯努利又名雅各布，诞生在瑞士的巴塞尔。他是伯努利家族的 8 个卓越的数学家之一(见 9.1 节伯努利数学世家族谱)。遵从他父亲的意愿，詹姆斯学习了神学并担任了神职。但和他父亲的愿望相反，他也研究数学和天文学。从 1676 年到 1682 年，他游历欧洲，获悉了数学和科学的最新发现。1682 年，他返回巴塞尔，创立了数学和科学学校。1687 年他被任命为巴塞尔大学的数学教授，并在这个位置终其一生。

詹姆斯·伯努利众所周知的著作是《推测术》(Ars Conjectandi)，在他死后 8 年得以发表。在这本著作中，他描述了在概率论和枚举中的已知结果，并常常对已知结果提供另外的证明。这本著作也包含了概率论对机会对策的应用和关于著名的大数定律的介绍。这条定律叙述了如果 $\varepsilon>0$，当 n 变得任意大时，事件 E 在 n 次试验中出现的次数除以 n 的比与 $p(E)$ 的差在 ε 之内的概率接近于 1。

显然应该如此。在这串等式中的第一处相等是二项式定理的结果，第二处相等是由于 $q=1-p$。

5.2.7 随机变量

许多问题都涉及一个与试验结果相关的数值。例如，我们可能想知道当随机产生 10 位二进位串时含 9 个 1 的概率，或者想知道掷 20 次硬币时有 11 次头像向下的概率。为了研究这类问题我们引入随机变量的概念。

定义 5 一个随机变量是从试验的样本空间到实数集的函数，即一个随机变量对每个可能的结果指派一个实数值。

注意 一个随机变量是一个函数，而不是一个变量，并且它也不是随机的！

例 10 假设一个硬币被掷 3 次。令 $X(t)$ 是出现头像的个数，其中 t 是结果。那么随机变量 $X(t)$ 取值如下：

$X(HHH)=3$

$X(HHT)=X(HTH)=X(THH)=2$

$X(TTH)=X(THT)=X(HTT)=1$

$X(TTT)=0$ ■

定义 6 一个随机变量 X 在样本空间 S 中的分布是对所有的 $r\in X(S)$ 的对 $(r, p(X=r))$ 的集合，其中 $p(X=r)$ 是 X 取值 r 的概率。通常是通过对每个 $r\in X(S)$ 指定 $p(X=r)$ 来描述分布的。

例 11 当掷 3 个硬币时，八种可能的结果中每一个出现的概率都是 1/8，例 10 中的随机变量 $X(t)$ 是由 $p(X=3)=1/8$，$p(X=2)=3/8$，$p(X=1)=3/8$，$p(X=0)=1/8$ 给出的。 ■

例 12 设 X 是掷一对骰子时出现的点数之和，那么这个随机变量怎样对 36 个可能的结果 (i, j) 取值？这里的 i 和 j 分别表示当掷两个骰子时，第一和第二个骰子出现的点数。

解 随机变量 X 取值如下：

$X((1, 1))=2$

$X((1, 2))=X((2, 1))=3$

$X((1, 3))=X((2, 2))=X((3, 1))=4$

$X((1, 4))=X((2, 3))=X((3, 2))=X((4, 1))=5$

$X((1, 5))=X((2, 4))=X((3, 3))=X((4, 2))=X((5, 1))=6$

$X((1, 6))=X((2, 5))=X((3, 4))=X((4, 3))=X((5, 2))=X((6, 1))=7$

$X((2, 6))=X((3, 5))=X((4, 4))=X((5, 3))=X((6, 2))=8$

$X((3, 6))=X((4, 5))=X((5, 4))=X((6, 3))=9$

$X((4, 6))=X((5, 5))=X((6, 4))=10$

$X((5, 6))=X((6, 5))=11$

$X((6, 6))=12$ ■

在 5.3 节将继续研究随机变量，我们也将说明它们在各种应用中是怎样使用的。

5.2.8 生日问题

有一个著名的问题：要使一个房间中至少有两个人同月同日生的可能性大于这种情况不存在的可能性，那么房间里至少需要多少人才行？大多数人发现这个答案（我们将在例 13 确定它）是惊人得小。在求解了这个著名的问题之后，我们将显示类似的推理是怎样适于求解与散列函数有关的问题的。

例 13　生日问题　如果要求房间中至少 2 个人有相同的生日的概率大于 1/2，那么所需的最少人数是多少？

解　首先叙述某些假设。我们假设房间中的人生于某一天是独立的。进一步，假设生于某一天是等可能的，并且一年是 366 天。（实际上一年的某些日子出生的人比其他日子更多，例如在新年这样的节日之后 9 个月的日子，此外只有闰年有 366 天。）

为了找到房间中的 n 个人里至少 2 个人生日相同的概率，首先计算这些人生日彼此都不相同的概率 p_n，那么至少 2 个人有同样的生日的概率是 $1-p_n$。为计算 p_n，我们考虑按照某个给定顺序的 n 个人的生日。想像他们一次一个人地进入房间，我们将计算每个即将进入房间的人与那些原来已经在房间的人有不同生日的概率。

第一个人与已经在房间的人的生日肯定不同。第二个人的生日与第一个人不同的概率为 365/366，这是因为第二个人除了诞生在第一个人的生日那天以外，出生在其余的 365 天的任何一天都有着不同的生日（这里和下面的步骤都用到某个人出生在一年的 366 天中的任何一天都是等可能的假设。）给定前两个人有不同生日的情况下，第三个人的生日与第一个人和第二个人两个人的生日都不相同的概率为 364/366，一般地，第 j 个人（$2\leqslant j\leqslant 366$）与已经进入房间的给定不同生日的 $j-1$ 个人有着不同生日的概率为

$$\frac{366-(j-1)}{366}=\frac{367-j}{366}$$

因为我们已经假设房间里的人的生日是独立的，我们可以断定房间的 n 个人有不同生日的概率是

$$p_n=\frac{365}{366}\frac{364}{366}\frac{363}{366}\cdots\frac{367-n}{366}$$

因此得到 n 个人中至少有两个人具有相同生日的概率是

$$1-p_n=1-\frac{365}{366}\frac{364}{366}\frac{363}{366}\cdots\frac{367-n}{366}$$

为确定使得其中至少两个人具有相同生日的概率大于 1/2 时房间里的最少人数，使用关于 $1-p_n$ 的公式，对于正在增长的 n 值进行计算，直到这个概率大于 1/2。（有一种使用微积分的更精确的方法可以省略这个计算。但是我们这里不用。）在经过可观的计算以后，我们发现对于 $n=22$，$1-p_n\approx 0.475$，而对于 $n=23$，$1-p_n\approx 0.506$. 因此，使得至少有两个人具有相同生日的概率大于 1/2 所需要的最少人数是 23。■

生日问题的解引出了例 14 中关于散列函数问题的解。

例 14　散列函数中碰撞的概率　回顾 2.4 节，一个散列函数 $h(k)$ 是从（存储在数据库中的记录的）关键字到存储地址的映射。散列函数把一个大范围的关键字（例如美国将近 3 亿个社会保险号）映射到小很多的存储地址的集合中。一个好的散列函数很少产生碰撞，所谓碰撞就是在一个给定应用中当相对较少的记录起作用时，两个不同的关键字映到相同的存储地址。对一个散列函数，没有两个关键字映到相同的地址，或者换句话说没有碰撞的概率是什么？

解　为计算这个概率，假设一个随机选择的关键字映到一个地址的概率是 $1/m$，其中 m 是有效地址的个数，即散列函数对关键字的分配是均匀的。（实际上散列函数不可能满足这个假设，但是对于好的散列函数这个假设近乎正确。）我们进一步假设被选中记录的关键字是关键字域上任何一个元素的概率相等，并且这些关键字都是独立选择的。

假设关键字是 k_1，k_2，…，k_n。当我们把第二个记录加入时，它被映到与第一个记录的地址不同的地址，即 $h(k_1)\neq h(k_2)$ 的概率是 $(m-1)/m$，因为第一个记录放入后还有 $m-1$ 个没用过的地址。在无碰撞放好第一个和第二个记录以后，第三个记录映到一个没有用过的地址的概率

是$(m-2)/m$。一般说来，前$j-1$个记录已经被无碰撞地映到地址$h(k_1)$，$h(k_2)$，…，$h(k_{j-1})$以后，第j个记录映到一个没有用过的地址的概率是$(m-(j-1))/m$，因为m个地址中已经用了$j-1$个地址。

由于关键字是独立的，所有n个关键字被映到不同地址的概率是

$$p_n=\frac{m-1}{m}\cdot\frac{m-2}{m}\cdot\cdots\cdot\frac{m-n+1}{m}$$

从而得到存在至少一个碰撞，即至少两个关键码被映到相同地址的概率是

$$1-p_n=1-\frac{m-1}{m}\cdot\frac{m-2}{m}\cdot\cdots\cdot\frac{m-n+1}{m}$$

我们可以使用微积分的技术找出对于给定的m值，使得碰撞概率大于一个特定值的最小的n值。可以证明使得碰撞概率大于1/2的最小的整数n近似是$n=1.177\sqrt{m}$。例如当$m=1\ 000\ 000$时，使得碰撞概率大于1/2的最小的整数n是1178。■

5.2.9 蒙特卡罗算法

本书至今讨论的算法都是确定型算法。即只要给定同样的输入，每个算法总是以同样的方式运行。但是有许多情况下我们希望算法在一步或多步中作随机选择。当一个确定型算法不得不遇到大量的、甚至是无数多种可能的情况时就会出现这种情况。在一步或者多步作随机选择的算法称为*概率算法*。这一节我们将讨论关于判定问题的一类特殊的概率算法，即*蒙特卡罗算法*。蒙特卡罗算法对于问题总能得到答案，但是这些答案可能存在很小的出错概率。当算法执行足够多的计算时，答案出错的概率迅速下降。判定问题以“真”或者“假”作为它们的答案。

求解判定问题的蒙特卡罗算法使用一系列的测试。只要测试执行得越多，算法正确回答判定问题的概率就增加。在算法的每一步，可能回答“真”，这意味着答案是“真”，不再需要进一步的迭代。回答“不知道”，这意味着答案可能是“真”也可能是“假”。这样一个算法经过多次迭代以后，如果任何一次迭代产生答案“真”，那么最后答案是“真”；如果每次迭代都产生答案“不知道”，那么答案就是“假”。如果正确的答案是“假”，算法回答“假”，因为每次迭代将产生“不知道”。但是，如果正确的答案是“真”，算法可能回答“真”或“假”，因为很可能每次迭代回答都是“不知道”，即使正确的答案是“真”。我们将证明当测试数增加时，这种可能性变得非常小。

给定答案为“真”，假设p是一次测试回答“真”的概率，从而$1-p$是给定答案为“真”回答“不知道”的概率。因为所有的n次迭代都产生答案“不知道”时算法才回答“假”，并且这些迭代执行独立的测试，因此出错的概率是$(1-p)^n$。当$p\neq0$时，这个概率随测试次数增加接近于0。所以，当答案是“真”时算法回答“真”的概率接近于1。

例15 质量控制 （这个例子来自[AhUl95]）假设一个工厂多批次定购处理器芯片，每批n片，其中n为正整数。芯片制造商只对这些批次中的某一批进行了测试以保证这批芯片中的所有芯片都是好的。在以前未测试的批次中，当做随机测试时，观察到一个特定芯片是坏芯片的概率是0.1。这家PC工厂想判定在一批芯片中是否所有的芯片都是好的。为此这家工厂可以对一批中的每片芯片进行测试，这需要测试n次。假设执行每次测试需要的时间为常数，这些测试需要$O(n)$秒。这家PC工厂能不能用更少的时间来检查这批芯片是否检测过？

解 只要我们能够接受一定的出错概率，就能用蒙特卡罗算法确定一批芯片是否已经被芯片制造商检测过。这个算法是为了回答下面的问题：“这批芯片是否被制造商测试过?”它一片接一片连续从这批中随机选择芯片并进行测试。当发现一片坏芯片时算法回答“真”并且停止。如果一个被测的芯片是好的，算法回答“不知道”并且继续下一个芯片。在算法测试了一定数量的芯片

以后，比如说k片，没有得到回答“真”，那么算法确定答案是“假”，即算法判定这批芯片是好的，也就是芯片制造商已经测试了这批芯片中的所有芯片。

这个算法回答出错只有一种情况，就是它断定一批没有测过的芯片已经被芯片制造商测过了。在一批没测过的芯片中一片芯片是好的概率是$1-0.1=0.9$。由于在一批芯片中测试不同芯片的事件是独立的，给定这批芯片没被测试过的情况下，算法的所有k步都产生答案“不知道”的概率是0.9^k。

取k足够大，我们就可以使得这个概率像我们所希望的那样小。例如，通过测试66片芯片，算法判定一批已经被芯片制造商测过的芯片的概率是0.9^{66}，它小于0.001。即算法回答出错的概率小于千分之一。注意到这个概率与这批芯片的片数n无关。这就是说，不管这批有多少片芯片，蒙特卡罗算法使用常数或者$O(1)$次测试，并且需要$O(1)$秒。只要PC工厂可以接受小于千分之一的错误率，蒙特卡罗算法将节省PC工厂很多的测试。如果需要更小的错误率，PC工厂在每批中可以测试更多的芯片；读者可以验证132次测试将使得错误率降低到1/1 000 000。 ■

例16 素数的概率测试 在第2章我们提到过一个合数，即一个不是素数的整数，对小于$n/4$的基b通过米勒测试，其中$1<b<n$(见2.6节练习30)。这个观察是蒙特卡罗算法确定一个正整数是否为素数的基础。因为大素数在公共密钥密码系统中起着基础的作用(见2.6节)，能够快速生成大素数就变得特别重要。

算法的目标是判定这个问题“n是合数吗?”给定一个正整数n，我们随机选择一个整数$b(1<b<n)$并且确定n能否通过基为b的米勒测试。如果n测试失败，因为n一定是合数，因此答案是“真”并且算法结束。否则，我们重复测试k次，其中k是整数。每次我们选择一个随机整数b，确定n能否通过对于基b的米勒测试。如果在每一步答案是“不知道”，算法回答“假”，即它说n不是合数，因而它是素数。当n是合数，并且在k步迭代的每一步输出答案是“不知道”，这是算法返回错误答案的唯一可能。对于一个随机选择的基b，一个合数n通过米勒测试的概率小于1/4。因为满足$1<b<n$的整数b在每次迭代时是随机选择的，并且这些迭代是独立的，所以n是合数但算法回答n是素数的概率小于$(1/4)^k$。通过取足够大的k，我们可以使得这个概率特别小。例如，用10次迭代，当n实际上是合数但算法判定n为素数的概率小于1/1 000 000。用30次迭代，这个概率小于$1/10^{18}$。这是一个特别小概率的事件。

为生成大素数，比如说200位。我们随机选择一个200位的整数n，并且运行这个算法，用30次迭代。如果算法判定n是素数，我们可以把它当成两个素数之一用在RSA密码系统的密钥中。如果n实际上是合数并且作为这个密钥的一部分来使用，这个用来解密信息的过程不会产生原始的被加密的信息。这个密钥将被抛弃而使用两个新的可能的素数。 ■

5.2.10 概率方法

第1章讨论了存在性证明并且说明在构造的存在性证明和非构造的存在性证明之间的区别。由保罗·厄多斯和阿尔弗雷德(任伊引入的概率方法是一个强有力的技术，它可以用于创建非构造的存在性证明。为了使用概率方法证明关于集合S的结果，例如S中一个具有指定性质元素的存在性，我们把概率赋给集合的元素。然后我们使用概率论的方法证明关于S的元素的结果。特别地，我们可以通过证明一个元素$x\in S$具有这个性质的概率是正的来证明具有特定性质的元素存在。这种概率方法是基于定理3的等价叙述。

定理3 *概率方法* 如果集合S的一个元素不具有一个特定性质的概率小于1，那么S中存在具有这条性质的元素。

一个基于概率方法的存在性证明是非构造性的，因为它未找出一个具有所要求性质的特定

元素。

我们通过找到一个关于拉姆赛数 $R(k, k)$ 的下界来说明概率算法的能力。回顾 4.2 节，$R(k, k)$ 等于晚会的最少人数，以保证其中至少有 k 个人两两都是朋友，或者至少有 k 个人两两都是敌人(假设任两个人要么是朋友要么是敌人)。

定理 4 如果 k 是一个整数，$k\geqslant 2$，那么 $R(k, k)\geqslant 2^{k/2}$。

证 注意到定理对于 $k=2$ 和 $k=3$ 成立，因为 $R(2, 2)=2$ 和 $R(3, 3)=6$，正如 4.2 节所显示的。现在假设 $k\geqslant 4$。我们将使用概率方法证明如果晚会上少于 $2^{k/2}$ 个人，可能没有 k 个人两两都是朋友，也没有 k 个人两两都是敌人。这就证明了 $R(k, k)$ 至少是 $2^{k/2}$。

为了使用概率方法，我们设两个人是朋友还是敌人是等可能的(注意这个假设不一定是真实的)。假设晚会上有 n 个人。因而晚会上有 $\binom{n}{k}$ 个不同的 k 个人的小组，记作 $S_1, S_2, \cdots, S_{\binom{n}{k}}$。设 E_i 是 S_i 中的所有 k 个人两两是朋友或两两是敌人的事件。在 n 个人中存在 k 个人两两是朋友或两两是敌人的概率等于 $p\left(\bigcup_{i=1}^{\binom{n}{k}} E_i\right)$。

根据我们的假设，两个人是朋友或是敌人是等可能的。两个人是朋友的概率等于他们是敌人的概率；两个概率都等于 1/2。进一步，因为 S_i 中有 k 个人，因此在 S_i 中存在 $\binom{k}{2}=k(k-1)/2$ 对人。于是，在 S_i 中所有的 k 个人两两都是朋友的概率与 S_i 中所有的 k 个人两两都是敌人的概率都等于 $(1/2)^{k(k-1)/2}$，从而得到 $p(E_i)=2(1/2)^{k(k-1)/2}$。

在一群 n 个人中存在 k 个人两两是朋友或两两是敌人的概率等于 $p\left(\bigcup_{i=1}^{\binom{n}{k}} E_i\right)$。使用布尔不等式(练习 15)，得到

$$p\left(\bigcup_{i=1}^{\binom{n}{k}} E_i\right) \leqslant \sum_{i=1}^{\binom{n}{k}} p(E_i) = \binom{n}{k}\cdot 2\left(\frac{1}{2}\right)^{k(k-1)/2}$$

根据 4.4 节练习 17，我们有 $\binom{n}{k}\leqslant n^k/2^{k-1}$，因此

$$\binom{n}{k}2\left(\frac{1}{2}\right)^{k(k-1)/2} \leqslant \frac{n^k}{2^{k-1}}2\left(\frac{1}{2}\right)^{k(k-1)/2}$$

现在如果 $n<2^{k/2}$，我们有

$$\frac{n^k}{2^{k-1}}2\left(\frac{1}{2}\right)^{k(k-1)/2} < \frac{2^{k(k/2)}}{2^{k-1}}2\left(\frac{1}{2}\right)^{k(k-1)/2} = 2^{2-(k/2)} \leqslant 1$$

最后一步的得出是由于 $k\geqslant 4$。

我们现在可以断言当 $k\geqslant 4$ 时，$p\left(\bigcup_{i=1}^{\binom{n}{k}} E_i\right)<1$。因此，这个补事件，即在这个晚会上不存在一组 k 个人两两都是朋友也不存在 k 个人两两都是敌人的概率大于 0。从而得到，如果 $n<2^{k/2}$，至少存在一个集合，使得它不包含 k 个人两两都是朋友或者两两都是敌人的子集。 □

练习

1. 当掷一个不均匀的硬币时，如果出现头像是不出现头像的可能性的3倍，那么出现头像的概率应该是多少？不出现头像的概率应该是多少？
2. 当掷一个被填充的骰子时，如果骰子出现3点这一面的可能性是其他五个面的2倍，求出每种结果的概率。
3. 当掷一个不均匀的骰子时，如果骰子出现2点或4点这一面的可能性是出现其他四个面的3倍，并且掷出一个2点或者一个4点是等可能的，求出每种结果的概率。
4. 当结果是等可能的时候，证明在拉普拉斯的概率定义下条件(i)和(ii)是满足的。
5. 一对骰子是被填充的，第一个骰子出现4点的概率是2/7，第二个骰子出现3点的概率是2/7，且每个骰子出现其他点数的概率是1/7。当掷出2个骰子时，点数之和为7的概率是多少？
6. 当我们随机选择{1, 2, 3}的一个排列时，这些事件的概率是什么？
   a)1在3前面。 b)3在1前面。 c)3在1前面且3在2前面。
7. 当我们随机选择{1, 2, 3, 4}的一个排列时，这些事件的概率是什么？
   a)1在4前面。 b)4在1前面。
   c)4在1前面且4在2前面。 d)4在1前面，4在2前面，且4在3前面。
   e)4在3前面且2在1前面。
8. 当我们随机选择{1, 2, …, n}的一个排列时，其中 $n \geqslant 4$，这些事件的概率是什么？
   a)1在2前面。 b)2在1前面。
   c)1紧挨着2前面。 d)n在1前面且$n-1$在2前面。
   e)n在1前面且n在2前面。
9. 当我们随机选择26个英语小写字母的一个排列时，这些事件的概率是什么？
   a)排列由按字母表相反顺序的字母组成。 b)z是排列的第一个字母。
   c)排列中z在a前面。 d)排列中a紧接在z的前面。
   e)排列中a紧接在m的前面，m紧接在z的前面。
   f)排列中m，n，o在字母表中的原来位置。
10. 当我们随机选择26个英语小写字母的一个排列时，这些事件的概率是什么？
    a)排列中的前13个字母按照字母表的顺序。
    b)a是排列的第一个字母且z是排列的最后一个字母。
    c)a和z在排列中彼此相邻。 d)a和b在排列中彼此不相邻。
    e)排列中a和z被至少23个字母分开。 f)排列中z在a和b两者的前面。
11. 假设E和F是事件，满足$p(E)=0.7$且$p(F)=0.5$。证明$p(E \cup F) \geqslant 0.7$和$p(E \cap F) \geqslant 0.2$。
12. 假设E和F是事件，满足$p(E)=0.8$且$p(F)=0.6$。证明$p(E \cup F) \geqslant 0.8$和$p(E \cap F) \geqslant 0.4$。
13. 证明如果E和F是事件，那么$p(E \cap F) \geqslant p(E)+p(F)-1$。这就是邦弗罗尼(Bonferroni)不等式。
14. 使用数学归纳法证明下述一般性的邦弗罗尼不等式：
    $$p(E_1 \cap E_2 \cap \cdots \cap E_n) \geqslant p(E_1)+p(E_2)+\cdots+p(E_n)-(n-1)$$
    其中E_1，E_2，…，E_n是n个事件。
15. ☞ 证明如果E_1，E_2，…，E_n是一个有限样本空间的事件，那么
    $$p(E_1 \cup E_2 \cup \cdots \cup E_n) \leqslant p(E_1)+p(E_2)+\cdots+p(E_n)$$
    这就是布尔不等式。
16. 如果E和F是独立的事件，证明$\overline{E}$和$\overline{F}$也是独立的事件。
17. ☞ 如果E和F是独立的事件，证明或反驳$\overline{E}$和F也必须是独立的事件。
18. a)两个人出生在一周的同一天的概率是多少？
    b)一组n个人里，至少2个人出生在一周的同一天的概率是什么？

c)要使得至少 2 个人出生在一周的同一天的概率大于 1/2，需要多少个人？

19. a)2 个人出生在一年的同一个月的概率是多少？

b)一组 n 个人里，至少 2 个人出生在一年的同一个月的概率是什么？

c)要使得至少 2 个人出生在一年的同一个月的概率大于 1/2，需要多少个人？

20. 要使得某个人的生日就在当天的概率大于 1/2，求出房间里最少的人数(假设生于每一天是等可能的并且这一年有 366 天)。

21. 要使得房间里的两个人都生在 4 月 1 日的概率大于 1/2，求出房间的最少人数(假设生于每一天是等可能的并且这一年有 366 天)。

*22. 只有闰年有 2 月 29 日。能被 4 整除但不能被 100 整除的都是闰年，能被 100 整除但不被 400 整除的不是闰年，能被 400 整除的是闰年。

a)应该用哪种关于生日的概率分布来反映 2 月 29 日出现次数的多少？

b)利用 a)的概率分布，一组 n 个人至少两个人有相同的生日的概率是什么？

23. 给定掷硬币第一次的头像在上，当一个均匀的硬币被掷 5 次时恰好 4 次头像在上的条件概率是什么？

24. 给定掷硬币第一次的头像在下，当一个均匀的硬币被掷 5 次时恰好 4 次头像在上的条件概率是什么？

25. 给定第一位是 1，随机产生 4 位二进制串并使得它至少包含 2 个连续的 0 的条件概率是什么？(假设是 1 和是 0 的概率相同。)

26. 随机产生 3 位二进位串，设 E 是这个串含有奇数个 1 的事件，F 是这个串以 1 开始的事件。E 和 F 是独立的吗？

27. 设 E 和 F 分别表示有 n 个孩子的家庭同时有男孩和女孩以及至多有 1 个男孩的事件。在下述每种条件下 E 和 F 是独立的吗？

a) $n=2$　　b) $n=4$　　c) $n=5$

28. 假定一个孩子是男孩的概率是 0.51，且生在一个家庭的孩子的性别是独立的。一个家庭有 5 个孩子，那么

a)恰有 3 个男孩的概率是什么？　　b)至少有 1 个男孩的概率是什么？

c)至少有 1 个女孩的概率是什么？　　d)所有的孩子性别都相同的概率是什么？

29. 一组 6 个人玩“单人出局”的游戏来确定谁买茶点。每个人掷一个均匀的硬币，如果一个人掷出的结果不和组中其他任何人相同，这个人就必须买茶点。在掷过一次硬币以后出现这种单人出局的概率是多少？

30. 10 位的二进位串，如果每位的产生是独立的，求出下列每种情况下随机产生不包含 0 的位串的概率：

a)一位为 0 和为 1 是等可能的。　　b)一位为 1 的概率是 0.6。

c)第 i 位为 1 的概率是 $1/2^i$，$i=1, 2, 3, \cdots, 10$。

31. 求有 5 个孩子的家庭没有男孩的概率，如果孩子的性别是独立的，且

a)是男孩和是女孩是等可能的。　　b)是男孩的概率是 0.51。

c)第 i 个孩子是男孩的概率是 $0.51-(i/100)$。

32. 10 位二进位串，如果每位的产生是独立的，分别求在 30 题的 a)，b)和 c)条件下随机产生一个以 1 开始或以 00 结尾的位串的概率。

33. 按照习题 31 的 a)，b)和 c)的条件，分别求出有 5 个孩子的家庭中第 1 个孩子是男孩或者最后 2 个孩子是女孩的概率。

34. 执行 n 次独立的伯努利试验，其中每次试验的成功概率为 p，求下述每种情况的概率。

a)没有 1 次成功的概率。　　b)至少 1 次成功的概率。

c)至多 1 次成功的概率。　　d)至少 2 次成功的概率。

35. 求出在下述每种情况下执行 n 次独立的伯努利试验时的概率，其中每次试验的成功概率为 p。

a)没有 1 次失败的概率。　　b)至少 1 次失败的概率。

c)至多 1 次失败的概率。　　d)至少 2 次失败的概率。

36. 使用数学归纳法证明：如果 $E_1, E_2, \cdots, E_n$ 是样本空间 S 中的 n 个两两不相交的事件的序列，其中 n 是正整数，那么 $p\left(\bigcup_{i=1}^{n} E_i\right)=\sum_{i=1}^{n} p(E_i)$。

*37. (要求微积分)证明如果 E_1，E_2，…是样本空间 S 中的两两不相交的事件的无限序列，那么 $p(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} p(E_i)$ 。[提示：利用练习 36 并且取极限。]

38. 在异地掷一对骰子，一个诚实的旁观者通知我们至少有一个骰子掷出 6 点。

a)以诚实的观察者提供的信息为条件，两个骰子掷出的点数之和等于 7 的概率是多少？

b)假设诚实的观察者告诉我们至少一个骰子是 5 点。以此为给定条件，两个骰子的点数之和等于 7 的概率是多少？

**39. 这个练习利用概率方法证明一个关于循环赛的结果。在一个具有 m 个游戏者的循环赛中，每两个人玩一个游戏，其中一个赢，另一个输。

我们想要找一个关于正整数 m 和 k 的条件($k<m$)，以使得这个循环赛的结果有下述性质的可能：对每 k 个游戏者的集合，存在 1 个游戏者赢了这个集合的每个成员。假设当 2 个游戏者竞争时每个游戏者赢得这个游戏是等可能的，并且假设不同游戏的结果是独立的，这使得我们可以使用概率推理得出关于循环赛的结论。设 k 是小于 m 的正整数，E 是对每个具有 k 个游戏者的集合 S，存在 1 个游戏者赢了 S 中所有 k 个人的事件。

a) 证明 $p(\overline{E}) \leqslant \sum_{j=1}^{\binom{m}{k}} p(F_j)$，这里的 F_j 是一个事件，表示在 $\binom{m}{k}$ 个 k 人集合构成的表中，没有 1 个游戏者赢了第 j 个集合的所有 k 个人。

b)证明 F_j 的概率是 $(1-2^{-k})^{m-k}$。

c)从(a)和(b)证明 $p(\overline{E}) \leqslant \binom{m}{k}(1-2^{-k})^{m-k}$，因此得知，如果 $\binom{m}{k}(1-2^{-k})^{m-k}<1$，一定存在一个竞赛具有所描述的性质。

d)使用(c)找出 m 的值，使得在 m 个游戏者的循环赛中对每 2 个游戏者的集合 S，都存在 1 个游戏者，赢了 S 中的 2 个人。对 3 个游戏者的集合重复这个问题。

*40. 设计一个蒙特卡罗算法以确定一个整数 1 到 n 的排列是已经被排序的(即按照递增顺序排列)还是随机的排列。如果确定这个表没有被排序，算法的某一步应该回答“真”，否则回答“不知道”。在 k 步以后，如果每步回答都是“不知道”，则算法判定这些整数是排好序的。证明随着步数的增加，算法产生一个错误结果的概率会极小。[提示：对每一步，测试某些元素是否在正确的顺序上。保证这些测试是独立的。]

41. 使用伪码写出在例 16 中描述的素数的概率测试算法。

5.3 期望值和方差

5.3.1 引言

一个随机变量的期望值，是对样本空间的所有元素，将元素的概率与这个随机变量在这个元素的值相乘之后求和。因此期望值是一个随机变量值的加权平均。一个随机变量的期望值为这个随机变量值的分布提供了一个中心点，我们使用随机变量期望值的概念可以求解许多问题，例如确定谁在赌博游戏中占有优势，也可以计算算法平均情形下的复杂度。一个随机变量的另一个有用的度量就是*方差*。它告诉我们这个随机变量的值分布得多么散。可以使用随机变量的方差帮助我们估计一个随机变量取那些远离它的期望值的概率。

5.3.2 期望值

许多问题可以用我们所期望的随机变量的取值，或者更精确地说是随机变量在大量试验中的平均值来表示。这类问题包含：当掷 100 次硬币时预期会出现多少次头像？在表中线性查找一个元素时预期的比较次数是多少？为研究这类问题，我们引入关于一个随机变量的期望值的概念。

定义 1 随机变量 $X(s)$在样本空间 S 的期望值(或期望)等于

$$E(X)=\sum_{s\in S}p(s)X(s)$$

注意当样本空间 S 有 n 个元素时，$S=\{x_1,x_2,\cdots,x_n\}$，$E(x)=\sum_{i=1}^{n}p(x_i)X(x_i)$

注意 这里我们仅关心与有限期望值相关的随机变量。

例 1 **骰子的期望值** 设 X 是掷一个骰子时出现的点数，什么是 X 的期望值?

解 随机变量 X 取值为 1，2，3，4，5 或 6，每个具有概率 1/6。从而得到

$$E(X)=\frac{1}{6}\cdot 1+\frac{1}{6}\cdot 2+\frac{1}{6}\cdot 3+\frac{1}{6}\cdot 4+\frac{1}{6}\cdot 5+\frac{1}{6}\cdot 6=\frac{21}{6}=\frac{7}{2}$$ ■

Extra Examples **例 2** 一个均匀的硬币被掷了 3 次，令 S 是 8 种可能结果的样本空间，X 是随机变量，它对结果的赋值是结果中的头像数。那么 X 的期望值是什么?

解 在 5.2 节例 10 我们列出了掷 3 次硬币时 X 对 8 个可能结果的值。由于硬币是均匀的且每次掷硬币是独立的，每个结果的概率都是 1/8。因此，

$$\begin{aligned}E(X)=&\frac{1}{8}[X(HHH)+X(HHT)+X(HTH)+X(THH)+X(TTH)\\&+X(THT)+X(HTT)+X(TTT)]\\=&\frac{1}{8}(3+2+2+2+1+1+1+0)=\frac{12}{8}\\=&\frac{3}{2}\end{aligned}$$

因此，当一个均匀的硬币被掷 3 次时出现头像的平均次数是 3/2。 ■

当一个试验有相对较少的结果时，我们可以直接从定义计算随机变量的期望值，正像在例 2 中所做的。但是，当一个试验有许多结果时，直接由定义计算随机变量的期望值可能是不方便的。我们可以换一种做法，把随机变量的值相等的试验结果分成组来找随机变量的期望值，正如定理 1 所示。

定理 1 如果 X 是随机变量，$p(X=r)$ 是 $X=r$ 的概率，即 $p(X=r)=\sum_{s\in S,X(s)=r}p(s)$，那么

$$E(X)=\sum_{r\in X(S)}p(X=r)r$$

证 假设 X 是域为 $X(S)$的随机变量，令 $p(X=r)$是随机变量 X 取值 r 的概率。因此，$p(X=r)$是使得 $X(s)=r$ 的结果 s 的概率之和。从而得到

$$E(X)=\sum_{r\in X(S)}p(X=r)r$$ □

例 3 和定理 2 的证明说明了这个公式的用法。在例 3 我们将找出当掷两个均匀的骰子时出现的点数之和的期望值。在定理 2 我们将找出当执行 n 次伯努利试验时，成功次数的期望值。

例 3 当掷一对均匀的骰子时所出现的点数之和的期望值是什么?

解 设 X 是随机变量，它等于掷一对骰子所出现的点数之和。在 5.2 节例 12，我们列出了关于这个试验的 36 个结果的 X 的值。X 的值域是{2，3，4，5，6，7，8，9，10，11，12}。由 5.2 节例 12 我们看到

$$p(X=2)=p(X=12)=1/36$$
$$p(X=3)=p(X=11)=2/36=1/18$$
$$p(X=4)=p(X=10)=3/36=1/12$$

$$p(X=5)=p(X=9)=4/36=1/9$$
$$p(X=6)=p(X=8)=5/36$$
$$p(X=7)=6/36=1/6$$

把这些值带入公式，得

$$\begin{aligned}E(X)&=2\cdot\frac{1}{36}+3\cdot\frac{1}{18}+4\cdot\frac{1}{12}+5\cdot\frac{1}{9}+6\cdot\frac{5}{36}+7\cdot\frac{1}{6}\\&\quad+8\cdot\frac{5}{36}+9\cdot\frac{1}{9}+10\cdot\frac{1}{12}+11\cdot\frac{1}{18}+12\cdot\frac{1}{36}\\&=7\end{aligned}$$

■

定理 2 当执行 n 次伯努利试验时，预期成功的次数是 np，这里 p 是每次试验成功的概率。

证 令 X 是等于 n 次试验中成功次数的随机变量。由 5.2 节定理 2，我们看到 $p(X=k)=C(n,k)p^kq^{n-k}$，于是有

$$\begin{aligned}E(X)&=\sum_{k=1}^{n}kp(X=k) &&\text{根据定理 1}\\&=\sum_{k=1}^{n}kC(n,k)p^kq^{n-k} &&\text{根据 5.2 节定理 2}\\&=\sum_{k=1}^{n}nC(n-1,k-1)p^kq^{n-k} &&\text{根据 4.4 节习题 21}\\&=np\sum_{k=1}^{n}C(n-1,k-1)p^{k-1}q^{n-k} &&\text{从每项中提出公因子 } np\\&=np\sum_{j=0}^{n-1}C(n-1,j)p^jq^{n-1-j} &&\text{使用 } j=k-1 \text{ 移动和式的下标}\\&=np(p+q)^{n-1} &&\text{根据二项式定理}\\&=np &&\text{由 } p+q=1\end{aligned}$$

因此，在 n 次伯努利试验中预期成功的次数是 np，证明完成。 □

5.3.3 期望的线性性质

定理 3 告诉我们期望值是线性的。例如，随机变量的和的期望值是它们的期望值之和。我们将发现这个性质特别有用。

定理 3 如果 $X_i(i=1,2,\cdots,n)$ 是 S 上的随机变量，n 是正整数，并且如果 a 和 b 是实数，那么

(i) $E(X_1+X_2+\cdots+X_n)=E(X_1)+E(X_2)+\cdots+E(X_n)$

(ii) $E(aX+b)=aE(X)+b$

证 对于 $n=2$ 第一个结果可以直接由期望值的定义得到，因为

$$\begin{aligned}E(X_1+X_2)&=\sum_{s\in S}p(s)(X_1(s)+X_2(s))\\&=\sum_{s\in S}p(s)X_1(s)+\sum_{s\in S}p(s)X_2(s)\\&=E(X_1)+E(X_2)\end{aligned}$$

使用数学归纳法，很容易从两个随机变量的情况得出具有 n 个随机变量的情况。最后，注意到 $E(aX+b)=\sum_{s\in S}p(s)(aX(s)+b)=a\sum_{s\in S}p(s)X(s)+b\sum_{s\in S}p(s)=aE(X)+b$，这是由于 $\sum_{s\in S}p(s)=1$。

□

例 4 和例 5 说明了怎样来使用定理 3。

例 4 用定理 3 求出掷一对均匀的骰子时所出现的点数之和的期望值(在例 3 中没有使用定理 3 也求出了这个值)。

解 设 X_1 和 X_2 是随机变量，其中 $X_1((i,j))=i$，$X_2((i,j))=j$，X_1 是第一个骰子上出现的点数，X_2 是第二个骰子上出现的点数。容易看出，因为$(1+2+3+4+5+6)/6=21/6=7/2$，所以 $E(X_1)=E(X_2)=7/2$。当掷两个骰子时，出现的两个点数之和就是和 X_1+X_2。根据定理 3，这个和的期望值是$E(X_1+X_2)=E(X_1)+E(X_2)=7/2+7/2=7$。 ■

例 5 在定理 2 的证明中，我们通过直接计算找到了执行 n 次伯努利试验时成功次数的期望值，其中 p 是每次试验成功的概率。说明怎样使用定理 3 找到这个结果。

解 设 X_i 是随机变量。如果 t_i 是成功，$X_i(t_1, t_2, \cdots, t_n)=1$；如果 t_i 是失败，$X_i(t_1, t_2, \cdots, t_n)=0$。$X_i$ 的期望值是 $E(X_i)=1\cdot p+0\cdot(1-p)=p$，$i=1, 2, \cdots, n$。令 $X=X_1+X_2+\cdots+X_n$ 使得 X 计数当执行 n 次伯努利试验时成功的次数。把定理 3 用于 n 个随机变量的和，就证明了 $E(X)=E(X_1)+E(X_2)+\cdots+E(X_n)=np$。 ■

我们利用期望的线性性质可以求解许多看起来很难的问题。要找一个随机变量的期望值，关键步骤就是把这个随机变量表示成一些很容易找到期望值的随机变量之和。例 6 和例 7 说明了这种技巧。

例 6 在帽子认领问题中的期望值 在一个餐厅里一个新雇员为 n 个人寄存帽子，他忘记在帽子上放寄存号。当顾客取帽子时这个寄存员随机选取留下的帽子交给他们。被正确返回的帽子数预期是多少?

解 设 X 是随机变量，它等于能够从寄存员那里取回自己帽子的人数。设 X_i 是随机变量，如果满足第 i 个人拿回自己的帽子则 $X_i=1$，否则 $X_i=0$。从而得到

$$X = X_1 + X_2 + \cdots + X_n$$

由于寄存员给这个人返回任何一顶帽子是等可能的，所以得出第 i 个人收到自己帽子的概率是 $1/n$。于是根据定理 1，对所有的 i，我们有

$$E(X_i) = 1\cdot p(X_i = 1) + 0\cdot p(X_i = 0) = 1\cdot 1/n + 0 = 1/n$$

根据期望的线性性质(定理 3)得到

$$E(X) = E(X_1) + E(X_2) + \cdots + E(X_n) = n\cdot 1/n = 1$$

于是，收到自己帽子的平均人数恰好是 1。注意这个答案与寄存帽子的人数是独立的!(我们将在 6.5 节找到对于没有一个人收到自己帽子的概率的显式公式。) ■

例 7 一个排列中逆序数的期望值 在前 n 个正整数的排列中，如果 $i<j$ 但是 j 在这个排列中位于 i 的前边，就称有序对(i, j)为排列的 1 个*逆序*。例如，在排列 3，5，1，4，2 中有 6 个逆序；这些逆序是：

$$(1,3), (1,5), (2,3), (2,4), (2,5), (4,5)$$

为了找出在前 n 个正整数的一个随机排列中期望的逆序数，我们令 $I_{i,j}$ 是前 n 个正整数的所有排列的集合上的随机变量，如果(i, j)是排列的逆序，则 $I_{i,j}=1$，否则为 0。这就得出，如果 X 是等于这个排列中逆序数的随机变量，那么

$$X = \sum_{1\leqslant i<j\leqslant n} I_{i,j}$$

注意到在一个随机选择的排列中 i 在 j 的前面还是 j 在 i 前面是等可能的(为此只要注意到具有每种性质的排列数相等就可以了)。于是，对于所有的对(i, j)，我们有

$$E(I_{i,j}) = 1\cdot p(I_{i,j} = 1) + 0\cdot p(I_{i,j} = 0) = 1\cdot 1/2 + 0 = 1/2$$

由于存在$\binom{n}{2}$个 i 和 j 的对$(1\leqslant i<j\leqslant n)$并且根据期望的线性性质(定理 3)我们有

$$E(X)=\sum_{1\leqslant i<j\leqslant n} I_{i,j}=\binom{n}{2}\cdot\frac{1}{2}=\frac{n(n-1)}{4}$$

从而得到在前 n 个正整数的一个排列中，平均存在 $n(n-1)/4$ 个逆序。 ■

5.3.4 平均情形下的计算复杂度

计算一个算法在平均情形下的计算复杂度可以转变成计算一个随机变量的期望值。设一个试验的样本空间是可能的输入 $a_j(j=1, 2, \cdots, n)$的集合，且令随机变量 X 对 a_j 的赋值是当 a_j 作为输入时该算法用到的操作数。基于我们对输入的了解，对每个可能的输入值 a_j 赋给一个概率 $p(a_j)$。那么该算法在平均情形下的复杂度是

$$E(X)=\sum_{j=1}^{n} p(a_j)X(a_j)$$

这就是 X 的期望值。

找一个算法平均情形下的计算复杂度通常比求它在最坏情形下的计算复杂度要困难得多，并且常常涉及复杂方法的使用。但是，也有一些算法，找出它在平均情形下的计算复杂度所需要的分析并不困难。例如，例 8 将说明怎样在不同的概率假设下找一个线性搜索算法的平均情形下的计算复杂度。这个概率是指我们搜索的元素是表中一个元素的概率。

例 8 线性搜索算法平均情形的复杂度 给定元素 x 和 n 个不同实数的表。在 2.1 节描述的线性搜索算法通过把这个元素与表中的每个元素进行比较来查找 x。当 x 被找到或者检查了所有的元素并确定 x 不在表中时算法结束。如果 x 在表中的概率是 p 并且 x 是表中 n 个元素的任一个都是等可能的，那么这个线性搜索算法在平均情形下的复杂度是什么？（存在 $n+1$ 种可能的输入。在表中的 n 个数与不在表中的 1 个数，这也作为 1 种单独的输入。）

解 在 2.3 节例 4 我们证明了如果 x 等于表中的第 i 个元素要用 $2i+1$ 次比较，在 2.3 节例 2 中又证明了如果 x 不在表中要用 $2n+2$ 次比较。x 等于表中第 i 个元素 a_i 的概率是 p/n，x 不在表中的概率是 $q=1-p$，从而得到线性搜索算法在平均情形下的计算复杂度是

$$\begin{aligned}E&=3p/n+5p/n+\cdots+(2n+1)p/n+(2n+2)q\\&=\frac{p}{n}(3+5+\cdots+(2n+1))+(2n+2)q\\&=\frac{p}{n}((n+1)^2-1)+(2n+2)q\\&=p(n+2)+(2n+2)q\end{aligned}$$

（第三个等式是由于 3.2 节的例 21。）例如，当 x 保证在表中时，有 $p=1$（对每个 i，$x=a_i$ 的概率是 $1/n$）和 $q=0$，因此 $E=n+2$，正如我们在 2.3 节例 4 所证明的。

当 x 在表中的概率 p 是 1/2 时，可知 $q=1-p=1/2$，从而 $E=(n+2)/2+n+1=(3n+4)/2$。类似地，如果 x 在表中的概率是 3/4，有 $p=3/4$ 和 $q=1/4$，因此 $E=3(n+2)/4+(n+1)/2=(5n+8)/4$。

最后，当 x 保证不在表中时，有 $p=0$ 和 $q=1$，从而得到 $E=2n+2$，这并不奇怪，因为我们必须搜索整个的表。 ■

例 9 说明了期望的线性性质可以帮助我们找到一个排序算法（即插入排序）的平均情形的复杂度。

例 9 插入排序的平均情形的复杂度 用插入排序对 n 个不同元素进行排序所使用的平均比较次数是多少？

解 首先假设 X 是随机变量，它等于用插入排序(如 2.1 节所述)对 n 个不同的元素的表 a_1，a_2，…，a_n 进行排序所用到的比较次数。那么 $E(X)$是使用的平均比较次数。(回顾对于 $i=2$，…，n，在第 i 步，插入排序将待排表中第 i 个元素插入由待排表前 $i-1$ 个元素已排好序的表中的适当位置。)

令 X_i 是随机变量，它等于在前 $i-1$ 个元素 a_1，a_2，…，a_{i-1}已经排序以后把 a_i 插入到合适位置使用的比较次数。由于

$$X = X_2 + X_3 + \cdots + X_n$$

我们可以使用期望的线性性质断定

$$E(X) = E(X_2 + X_3 + \cdots + X_n) = E(X_2) + E(X_3) + \cdots + E(X_n)$$

为了求出 $E(X_i)$，$i=2, 3, \cdots, n$，令 $p_j(k)$表示在这个表的前 j 个元素中的最大元素出现在第 k 个位置的概率，即 $\max(a_1, a_2, \cdots, a_j)=a_k$ 的概率，其中 $1\leqslant k\leqslant j$。由于表的元素是随机分布的，前 j 个元素中的最大元素出现在任何位置是等可能的。因此，$p_j(k)=1/j$。一旦 a_1，a_2，…，a_{i-1}已经被排序，如果 $X_i(k)$等于用插入排序将 a_i 插入到表中的第 k 个位置所用的比较次数，那么得到 $X_i(k)=k$。由于 a_i 可能插入到表的前 i 个位置中的任何一个位置，我们得到

$$E(X_i) = \sum_{k=1}^{i} p_i(k) \cdot X_i(k) = \sum_{k=1}^{i} \frac{1}{i} \cdot k = \frac{1}{i} \cdot \sum_{k=1}^{i} k = \frac{1}{i} \cdot \frac{i(i+1)}{2} = \frac{i+1}{2}$$

从而得到

$$\begin{aligned} E(X) &= \sum_{i=2}^{n} E(X_i) = \sum_{i=2}^{n} \frac{i+1}{2} = \frac{1}{2} \sum_{j=3}^{n+1} j \\ &= \frac{1}{2} \frac{(n+1)(n+2)}{2} - \frac{1}{2}(1+2) = \frac{n^2+3n-4}{4} \end{aligned}$$

为得到第三个等式，置 $j=i+1$ 来对和式的下标进行移位。为得到第四个等式，使用了公式 $\sum_{k=1}^{m} k=m(m+1)/2$(来自 3.2 节表 3-1)，其中 $m=n+1$，同时从中减去 $j=1$ 和 $j=2$ 这些缺失的项。我们得出结论，由插入排序对 n 个元素进行排序使用的平均比较次数等于$(n^2+3n-4)/4$，这是 $\Theta(n^2)$。 ■

5.3.5 几何分布

考虑例 10 中的问题。

例 10 设掷一个硬币出现头像向下的概率是 p，重复掷这个硬币直到头像向下为止。请问预期要掷多少次？

解 我们首先注意到样本空间由所有以任何个数的头像向上作为开始后跟一个头像向下的序列所组成。将头像向上记为 H，头像向下记为 T。那么样本空间是集合$\{T, HT, HHT, HHHT, HHHHT, \cdots\}$。注意这是一个无穷样本空间。看到掷硬币是独立的并且出现头像向上的概率为 $1-p$，我们就可以确定样本空间一个元素的概率。于是，$p(T)=p$，$p(HT)=(1-p)p$，$p(HHT)=(1-p)^2p$，一般说来，掷了 n 次硬币出现了头像向下，即在 $n-1$ 个头像向上出现之后跟随着一个头像向下的概率是$(1-p)^{n-1}p$。(练习 14 要求验证样本空间中的点的概率之和是 1。)

现在令 X 是随机变量，它等于在样本空间中的一个元素中掷硬币的次数。即 $X(T)=1$，$X(HT)=2$，$X(HHT)=3$，等等。注意 $p(X=j)=(1-p)^{j-1}p$，直到硬币出现头像向下为止。掷硬币的预期次数等于 $E(X)$。

依据定理 1，我们发现

$$E(X)=\sum_{j=1}^{\infty}j\cdot p(X=j)=\sum_{j=1}^{\infty}j(1-p)^{j-1}p=p\sum_{j=1}^{\infty}j(1-p)^{j-1}=p\cdot\frac{1}{p^2}=\frac{1}{p}$$

（上述推导中第三个等式根据 3.2 节表 3-2，$\sum_{j=1}^{\infty}j(1-p)^{j-1}=1/(1-(1-p))^2=1/p^2$ 。）于是，直到硬币出现头像向下为止，掷硬币的预期次数是 $1/p$。注意当硬币是均匀的时我们有 $p=1/2$，因此直到硬币出现头像向下为止，掷硬币的预期次数是 $1/(1/2)=2$。 ■

与掷硬币直到出现头像向下为止的预期次数相等的随机变量 X 是一个具有几何分布的随机变量的实例。

定义 2 如果对于 $k=1, 2, 3, \cdots$，$p(X=k)=(1-p)^{k-1}p$，那么随机变量 X 具有带参数 p 的几何分布。

由于几何分布用于研究在一个特定事件发生前所需要的时间，因而出现在许多应用中。例如在我们找到一个具有确定性质的物体之前需要的时间，在一个试验成功之前尝试的次数，又如一个产品在它失效之前可以使用的次数，等等。

当我们计算在硬币头像向下之前所要掷的次数的期望值的时候，就证明了定理 4。

定理 4 如果随机变量 X 有着带参数 p 的几何分布，那么 $E(X)=1/p$。

5.3.6 独立随机变量

我们已经讨论了独立的事件，现在将定义两个独立的随机变量意味着什么。

定义 3 随机变量 X 和 Y 在样本空间 S 上是独立的，如果

$$p(X(s)=r_1 \quad 且 \quad Y(s)=r_2)=p(X(s)=r_1)\cdot p(Y(s)=r_2)$$

换句话说，对一切实数 r_1 和 r_2，$X(s)=r_1$ 且 $Y(s)=r_2$ 的概率等于 $X(s)=r_1$ 的概率与 $Y(s)=r_2$ 的概率之积。

例 11 例 4 的随机变量 X_1 和 X_2 是独立的吗？

解 设 $S=\{1, 2, 3, 4, 5, 6\}$，i，j 属于 S。由于掷一对骰子有 36 个可能的结果并且每个结果是等可能的，故

$$p(X_1=i 且 X_2=j)=1/36$$

又由于第一个骰子出现 i 和第二个骰子出现 j 的概率都是 1/6，即 $p(X_1=i)=1/6$ 且 $p(X_2=j)=1/6$，从而有

$$p(X_1=i 且 X_2=j)=1/36=(1/6)(1/6)=p(X_1=i)p(X_2=j)$$

因此 X_1 和 X_2 是独立的。 ■

例 12 证明随机变量 X_1 和 $X=X_1+X_2$ 不是独立的，其中 X_1 和 X_2 的定义在例 4 中给出。

解 因为 $X_1=1$ 的含义是第一个骰子出现点数为 1，这就推出两个骰子的点数之和不可能等于 12，所以 $p(X_1=1 且 X=12)=0$。另一方面，$p(X_1=1)=1/6$ 和 $p(X=12)=1/36$。因此，$p(X_1=1 且 X=12)\neq p(X_1=1)\cdot p(X=12)$。这个反例证明了 X_1 和 X 不是独立的。 ■

两个独立的随机变量之积的期望值是它们的期望值之积，正如定理 5 所述。

定理 5 如果 X 和 Y 是样本空间 S 上的独立的随机变量，那么 $E(XY)=E(X)E(Y)$。

证 因为 X 和 Y 是独立的随机变量，由期望值的定义得到

$$\begin{aligned}E(XY)&=\sum_{s\in S}X(s)Y(s)p(s)\\&=\sum_{r_1\in X(S),r_2\in Y(S)}r_1r_2\cdot p(X(s)=r_1 且 Y(s)=r_2)\end{aligned}$$

$$= \sum_{r_1 \in X(S), r_2 \in Y(S)} r_1 r_2 \cdot p(X(s) = r_1) \cdot p(Y(s) = r_2)$$

$$= \left(\sum_{r_1 \in X(S)} r_1 p(X(s) = r_1) \right) \cdot \left(\sum_{r_2 \in Y(S)} r_2 p(Y(s) = r_2) \right)$$

$$= E(X)E(Y)$$

定理得证。 □

注意当 X 和 Y 不是独立的随机变量时，我们不能断定 $E(XY)=E(X)E(Y)$，正如例 13 所示。

例 13 设 X 和 Y 是计数一个硬币掷两次时出现头像和不出现头像的次数。由于 $p(X=2)=1/4$，$p(X=1)=1/2$，$p(X=0)=1/4$，根据定理 1，我们有

$$E(X) = 2 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{4} = 1$$

类似的计算显示 $E(Y)=1$。我们注意到出现两次头像向上且没有头像向下或者两次头像向下且没有头像向上时 $XY=0$，并且当出现一次头像向上和一次头像向下时 $XY=1$。因此

$$E(XY) = 1 \cdot \frac{1}{2} + 0 \cdot \frac{1}{2} = \frac{1}{2}$$

从而得到

$$E(XY) \neq E(X)E(Y)$$

这与定理 5 不矛盾，因为 X 和 Y 不独立，这一点读者应该可以验证(见练习 16)。 ■

5.3.7 方差

一个随机变量的期望值告诉我们的是它的平均值，但是并没有说明它的值的分布范围。例如，如果 X 和 Y 是集合 $S=\{1, 2, 3, 4, 5, 6\}$ 上的随机变量，对所有的 $s \in S$ 有 $X(s)=0$，且若 $s \in \{1, 2, 3\}$ 则 $Y(s)=-1$；若 $s \in \{4, 5, 6\}$ 则 $Y(s)=1$。那么 X 和 Y 的期望值都是 0。但是随机变量 X 永远等于 0，而随机变量 Y 总是与 0 相差 1。一个随机变量的方差帮助我们刻划一个随机变量的值的分布范围。

定义 4 设 X 是样本空间 S 上的随机变量。X 的方差记为 $V(X)$，且

$$V(X) = \sum_{s \in S} (X(s) - E(X))^2 p(s)$$

X 的标准差定义为 $\sqrt{V(X)}$，记作 $\sigma(X)$。

定理 6 提供了关于随机变量的方差的一个有用的简单表达式。

定理 6 如果 X 是样本空间 S 上的随机变量，那么 $V(X)=E(X^2)-E(X)^2$。

证 注意到

$$V(X) = \sum_{s \in S} (X(s) - E(X))^2 p(s)$$

$$= \sum_{s \in S} X(s)^2 p(s) - 2E(X) \sum_{s \in S} X(s) p(s) + E(X)^2 \sum_{s \in S} p(s)$$

$$= E(X^2) - 2E(X)E(X) + E(X)^2$$

$$= E(X^2) - E(X)^2$$

在倒数第二步我们使用了 $\sum_{s \in S} p(s) = 1$ 这一事实。 □

例 14 一个伯努利试验成功则 $X(t)=1$，失败则 $X(t)=0$。如果 p 是成功的概率，那么随机变量 X 的方差是什么?

解 因为 X 取值只能为 0 和 1，因此 $X^2(t)=X(t)$。于是，

$$V(X) = E(X^2) - E(X)^2 = p - p^2 = p(1-p) = pq$$ ■

例 15 一个骰子的值的方差 X 是掷一个骰子时出现的点数，什么是随机变量 X 的方差？

解 我们有 $V(X)=E(X^2)-E(X)^2$。由例 1 我们知道 $E(X)=7/2$。为求 $E(X^2)$，注意到 X^2 取值 i^2，$i=1, 2, \cdots, 6$。每个具有概率 1/6。从而得到

$$E(X^2) = \frac{1}{6}(1^2 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) = \frac{91}{6}$$

于是有

$$V(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$ ■

例 16 随机变量 $X((i, j))=2i$ 的方差是什么？这里的 i 和 j 是掷两个骰子时第一个骰子和第二个骰子上出现的点数。

解 我们将使用定理 6 找出 X 的方差。为此，我们需要找到 X 和 X^2 的期望值。注意到当 $k=2, 4, 6, 8, 10, 12$ 时 $p(X=k)$是 1/6，否则为 0。因而有

$$E(X) = (2+4+6+8+10+12)/6 = 7$$

和

$$E(X^2) = (2^2+4^2+6^2+8^2+10^2+12^2)/6 = 182/3$$

由定理 6 得

$$V(X) = E(X^2) - E(X)^2 = 182/3 - 49 = 35/3$$ ■

另一个有用的关于方差的事实是，两个独立的随机变量的和的方差是它们的方差之和。例如，这个结果可用于计算 n 个独立的伯努利试验结果的方差。

定理 7 如果 X 和 Y 是样本空间 S 上两个独立的随机变量，那么 $V(X+Y)=V(X)+V(Y)$。此外，如果对于正整数 n，X_i 是 S 上两两独立的随机变量，$i=1, 2, \cdots, n$，那么$V(X_1+X_2+\cdots+X_n)=V(X_1)+V(X_2)+\cdots+V(X_n)$。

证 由定理 6 有

$$V(X+Y) = E((X+Y)^2) - E(X+Y)^2$$

从而有

$$\begin{aligned} V(X+Y) &= E(X^2+2XY+Y^2) - (E(X)+E(Y))^2 \\ &= E(X^2) + 2E(XY) + E(Y^2) - E(X)^2 - 2E(X)E(Y) - E(Y)^2 \end{aligned}$$

因为 X 和 Y 是独立的，由定理 5 我们有 $E(XY)=E(X)E(Y)$，从而得到

$$\begin{aligned} V(X+Y) &= (E(X^2) - E(X)^2) + (E(Y^2) - E(Y)^2) \\ &= V(X) + V(Y) \end{aligned}$$

具有 n 个两两独立的随机变量的情况可以使用数学归纳法证明；这个证明留给读者完成。□

例 17 设掷两个骰子时随机变量 X 的值是 $X((i, j))=i+j$，其中 i 是第一个骰子出现的点数，j 是第二个骰子出现的点数。求 X 的方差和标准差。

解 设 X_1 和 X_2 是掷骰子的随机变量，其中 $X_1((i, j))=i$，$X_2((i, j))=j$。那么正如例 11 证明的，$X=X_1+X_2$ 和 X_1 与 X_2 都是独立的。由定理 7 得到 $V(X)=V(X_1)+V(X_2)$。与例 16 类似的简单计算与本章后的补充练习 27 告诉我们，$V(X_1)=V(X_2)=35/12$。因此，$V(X)=35/12+35/12=35/6$ 且 $\sigma(X)=\sqrt{35/6}$。■

我们现在求随机变量的方差，该随机变量计数了执行 n 次独立的伯努利试验时的成功次数。

例 18 当执行 n 次独立的伯努利试验时，什么是计数成功次数的随机变量的方差？这里 p

是每次试验成功的概率。

解　设 X_i 是随机变量，且若 t_i 是成功，则 $X_i((t_1, t_2, \cdots, t_n))=1$，若 t_i 是失败，则 $X_i((t_1, t_2, \cdots, t_n))=0$。令 $X=X_1+X_2+\cdots+X_n$，那么 X 计数在 n 次试验中的成功次数。由定理 7 得到 $V(X)=V(X_1)+V(X_2)+\cdots+V(X_n)$。使用例 14，有 $V(X_i)=pq$，$i=1, 2, \cdots, n$。从而得到 $V(X)=npq$。■

5.3.8 切比雪夫不等式

一个随机变量的取值与它的期望值差多少？下面的定理叫做切比雪夫[⊖]不等式，它对随机变量的值与它的期望值之差超过某个指定量的概率提供了一个上界，有助于回答这个问题。

定理 8　切比雪夫不等式　设 X 是在样本空间 S 上的概率函数为 p 的随机变量。如果 r 是一个正实数，那么

$$p(|X(s)-E(X)|\geqslant r)\leqslant V(X)/r^2$$

证　设 A 是事件

$$A=\{s\in S \mid |X(s)-E(X)|\geqslant r\}$$

我们想要证明的是 $p(A)\leqslant V(X)/r^2$。注意

$$\begin{aligned} V(X) &= \sum_{s\in S}(X(s)-E(X))^2p(s) \\ &= \sum_{s\in A}(X(s)-E(X))^2p(s)+\sum_{s\notin A}(X(s)-E(X))^2p(s) \end{aligned}$$

在这个表达式中的第二个和是非负的，因为它的每个被加数是非负的。又因为对于 A 中的每个元素 s，有 $(X(s)-E(X))^2\geqslant r^2$，这个表达式的第一个和至少是 $\sum_{s\in A}r^2p(s)$，因此，$V(x)\geqslant\sum_{s\in A}r^2p(s)=r^2p(A)$，这正是我们想证明的。□

例 19　当计数头像向下时与平均值的偏差　假设 X 是当一个均匀的硬币被掷 n 次时计数头像向下次数的随机变量。注意 X 是执行 n 次独立的伯努利试验时成功的次数，每次成功的概率是 1/2。因此得到 $E(X)=n/2$（根据定理 2）和 $V(X)=n/4$（根据例 18）。令 $r=\sqrt{n}$，使用切比雪夫定理得到

$$p(|X(s)-n/2|\geqslant\sqrt{n})\leqslant(n/4)/(\sqrt{n})^2=1/4$$

所以当一个均匀的硬币被掷 n 次时头像向下的次数与平均值的偏差大于 $\sqrt{n}$ 的概率不大于 1/4。■

切比雪夫不等式尽管可以用于任何随机变量，但在实际估计一个随机变量的值大大超过它的平均值的概率时常常失效。这可以用下面的例子说明。

例 20　设 X 是当掷一个均匀骰子时的随机变量，X 的值就是出现的点数。我们有 $E(X)=7/2$（见例 1）和 $V(X)=35/12$（见例 15）。因为 X 的可能取值是 1，2，3，4，5 和 6，$E(X)=7/2$，



⊖ **切比雪夫（Pafnuty Lvovich Chebyshev，1821—1894）**　切比雪夫出生于一个俄罗斯名门。他的父亲是一名退役军官，曾与拿破仑作战。在 1832 年，这个家庭带着 9 个孩子搬到莫斯科。在那里，切比雪夫修完了高中课程，并进入莫斯科大学的物理数学系学习。还是一个学生时，他就提出了一种新的求方程的近似根的方法。1841 年他从莫斯科大学毕业，获得数学学位，并且继续学习，在 1843 年通过硕士考试并在 1846 年完成硕士论文。

1847 年，切比雪夫聘为圣彼得堡大学的助教。在 1847 年他通过论文答辩。在 1860 年，他成为圣彼得堡大学的教授，并一直工作到 1882 年。他在 1849 年写的有关同余理论的著作对数论的发展影响很大。他关于素数分布的研究工作也是开创性的。他证明了贝川（Bertrand）的猜想，即对每个整数 $n>3$，存在一个在 n 和 $2n-2$ 之间的素数。切比雪夫提出了一些新的思想，后来用这些思想证明了素数定理。切比雪夫用多项式做函数逼近，这被广泛地用于计算机中对函数的求值。切比雪夫对力学也很感兴趣。他研究了怎样通过力偶将旋转运动转换成直线运动。切比雪夫平动是用三个连接在一起的棒体来逼近直线运动。

X 不可能比它的平均值多 5/2。因此，如果 $r>5/2$，$p(|X-7/2|\geqslant r)=0$。由切比雪夫不等式知道 $p(|X-7/2|\geqslant r)\leqslant(35/12)/r^2$。例如，当 $r=3$ 时，切比雪夫不等式告诉我们 $p(|X-7/2|\geqslant 3)\leqslant(35/12)/9=35/108$，这是一个很差的估计，因为 $p(|X-7/2|\geqslant 3)=0$。 ■

练习

1. 当一个均匀的硬币被掷 5 次时，头像在上的预期次数是多少？
2. 当一个均匀的硬币被掷 10 次时，头像在上的预期次数是多少？
3. 当一个均匀的骰子被掷 10 次时，出现 6 点的预期次数是多少？
4. 一个硬币是不均匀的，使得掷出头像的概率是 0.6，当掷 10 次时，头像在上的预期次数是多少？
5. 掷 2 个不均匀的骰子，其中出现 3 点的次数是出现其他每个点数的 2 倍。2 个骰子预期出现的点数和是什么？
6. 如果彩票包含了从集合{1，2，3，…，50}选出的 6 个中奖数字就赢奖 1000 万美元，否则不中奖，那么买 1 美元彩票中奖的期望值是多少？
7. 离散数学课程的期末考试有 50 道真假判断题，每道题 2 分；还有 25 道多选题，每道题 4 分。琳达正确回答判断题的概率是 0.9，正确回答多选题的概率是 0.8。她在期末考试预期的分数是多少？
8. 当掷 3 个均匀的骰子时预期出现的数字和是多少？
9. 表中含有 n 个不同的整数，假设 x 在这个表中的概率是 2/3，且 x 等于表中任何元素是等可能的。求由线性搜索算法找到 x 或确定它不在表中所用的平均比较次数。
10. 假设我们掷一个硬币直到它出现 2 次头像在下或者已经掷了 6 次为止，我们掷硬币的预期次数是多少？
11. 假设我们掷一个骰子直到出现 6 点或者已经掷了 10 次为止，我们掷骰子的预期次数是多少？
12. 假设掷一个骰子直到出现 6 点为止。
    a)我们掷 n 次骰子的概率是多少？　　b)我们掷骰子的预期次数是多少？
13. 假设我们掷一对骰子直到其点数之和是 7 为止。我们掷骰子的预期次数是多少？
14. 证明具有带参数 $p(0<p\leqslant 1)$ 的几何分布的随机变量的概率之和等于 1。
15. 证明如果随机变量 X 有带参数 p 的几何分布，且 j 是正整数，那么 $p(X\geqslant j)=(1-p)^{j-1}$。
16. 设 X 和 Y 是当掷两个硬币时计数出现头像在上和头像在下次数的随机变量。证明 X 和 Y 不是独立的。
17. 如果 1000 位的整数是素数的概率近似是 1/2302，估计需要随机选择 1000 位整数以找到一个素数所需要的预期次数。
18. 设 X 和 Y 是随机变量，并且对于样本空间 S 的所有的点，X 和 Y 是非负的。设 Z 是如下定义的随机变量：对所有的元素 $s\in S$，$Z(s)=\max(X(s),Y(s))$。证明 $E(Z)\leqslant E(X)+E(Y)$。
19. 掷两个骰子，设 X 是出现在第一个骰子上的点数，Y 是出现在两个骰子上的点数之和。证明 $E(X)E(Y)\neq E(XY)$。
20. 设 A 是事件，I_A 是 A 的指示器随机变量，如果 A 出现则 I_A 等于 1，否则为 0。证明 A 的指示随机变量的期望等于 A 的概率，即 $E(I_A)=p(A)$。
21. 一个*顺串*是在一个伯努利试验序列中的极大的成功序列。例如，在序列 S,S,S,F,S,S,F,F,S 中，其中 S 代表成功，F 代表失败，这里存在三个顺串，分别由 3 个成功，2 个成功，1 个成功组成。设 R 是 n 次独立伯努利试验的序列集合上的随机变量，它计数了在这个序列中的顺串的个数。求 $E(R)$。
    [提示：证明 $R=\sum_{j=1}^{n} I_j$，如果一个顺串在第 j 次伯努利试验开始，则 $I_j=1$，否则为 0。找到 $E(I_1)$，然后求 $E(I_j)$，其中 $1<j\leqslant n$。]
22. 设 $X(s)$是随机变量，对所有 $s\in S$，$X(s)$是非负整数，且 A_k 是满足 $X(s)\geqslant k$ 的事件。证明
$$E(X)=\sum_{k=1}^{\infty}p(A_k)$$
23. 当一个均匀的硬币被掷 10 次时，头像在上次数的方差是什么？
24. 当一个均匀的骰子被掷 10 次时，出现 6 点的次数的方差是什么？

25. 设 X_n 是掷 n 个硬币时计数头像在下次数和头像在上次数之差的随机变量。
    a) X_n 的期望值是什么？ b) X_n 的方差是什么？
26. 提供一个例子说明当两个随机变量不独立时，它们的和的方差不一定等于它们的方差之和。
27. 使用切比雪夫不等式找出一个均匀的硬币被掷 n 次时，出现头像在下的次数与平均值的偏差超过 $5\sqrt{n}$ 的概率的上界。
28. 掷一个不均匀的硬币，其中出现头像的概率等于 0.6。使用切比雪夫不等式找出掷这个硬币 n 次时，出现头像在下的次数与平均值的偏差超过$\sqrt{n}$的概率的上界。
29. 设 X 为样本空间 S 上的随机变量，且对于所有的 $s\in S$，$X(s)\geqslant 0$。证明对每个正实数 a，$p(X(s)\geqslant a)\leqslant E(X)/a$。这个不等式叫做马尔可夫不等式。
30. 假设一个灌装厂一天灌装苏打饮料的听数是一个随机变量。它的期望值是 10 000，方差是 1000。
    a)使用马尔可夫不等式(练习 29)得到该厂在某一天灌装听数超过 11 000 的概率的上界。
    b)使用切比雪夫不等式得到该厂在某一天灌装听数在 9000 到 11 000 之间的概率的下界。
31. 假设一个回收中心一天回收的罐头盒数是一个随机变量，它的期望值是 50 000，方差是 2500。
    a)使用马尔可夫不等式(练习 29)得到该中心在某一天回收罐头盒数超过 55 000 的概率的上界。
    b)使用切比雪夫不等式提供该中心在某一天回收的罐头盒数在 40 000 到 60 000 之间的概率的下界。

*32. 设 x 是 n 个不同整数的表中第 i 个数的概率为 $i/[n(n+1)]$。通过线性搜索算法找到 x 或者确定 x 不在表中，求该算法使用的平均比较次数。

*33. 在这个练习中我们要导出对一个变种的冒泡排序算法的平均情形下复杂度的估计。这个算法一旦做了一次没有交换的扫描就结束。设 $\{a_1, a_2, \cdots, a_n\}$ 是 n 个不同整数的集合，$a_1<a_2<\cdots<a_n$，X 是该集合的排列的集合上的随机变量，且 $X(P)$ 等于通过这个冒泡排序将排列 P 中的整数排成递增顺序时所用的比较次数。
    a)在输入对这些整数的 $n!$ 个排列都是等可能的假设下，证明这个冒泡排序用到的平均比较次数等于$E(X)$。
    b)使用 2.3 节例 5 证明 $E(X)\leqslant n(n-1)/2$。
    c)证明这个排序对输入中每两个整数的逆序至少做一次比较。
    d)设 $I(P)$ 是随机变量，它等于排列 P 中的逆序数。证明 $E(X)\geqslant E(I)$。
    e)设 $I_{j,k}$ 是随机变量，如果在排列 P 中 a_k 在 a_j 前面 $I_{j,k}(P)=1$，否则 $I_{j,k}=0$。证明 $I(P)=\sum_{k}\sum_{j<k} I_{j,k}(P)$。
    f)证明 $E(I)=\sum_{k}\sum_{j<k} E(I_{j,k})$。
    g)证明 $E(I_{j,k})=1/2$。[提示：证明 $E(I_{j,k})$ = 在排列 P 中 a_k 出现在 a_j 前面的概率。然后证明在一个排列中 a_k 出现在 a_j 前面和 a_j 出现在 a_k 前面是等可能的。]
    h)使用(f)和(g)证明 $E(I)=n(n-1)/4$。
    i)从(a)、(b)和(h)得出排序 n 个整数使用的平均比较次数是 $\Theta(n^2)$ 的结论。

*34. 在这个练习中我们找在 3.5 节练习 38 的导言中描述的快速排序算法的平均情形下的复杂度，这里假设在排列的集合上是平均分布的。
    a)设 X 是快速排序算法对 n 个不同整数的表排序用的比较次数。证明快速排序算法使用的平均比较次数是 $E(X)$(这里的样本空间是 n 个整数的所有 $n!$ 个排列的集合)。
    b)设 $I_{j,k}$ 表示随机变量，如果初始数表的第 j 个最小元素和第 k 个最小的元素在快速排序算法排序这个数表时曾经被比较过，则它等于 1，否则为 0。证明 $X=\sum_{k=2}^{n}\sum_{j=1}^{k-1} I_{j,k}$。
    c)证明 $E(X)=\sum_{k=2}^{n}\sum_{j=1}^{k-1} p$(第 j 个最小元素和第 k 个最小元素被比较过的概率)。
    d)证明 p(第 j 个最小元素和第 k 个最小元素被比较过的概率)等于 $2/(k-j+1)$，其中 $k>j$。
    e)使用(c)和(d)证明 $E(X)=2(n+1)\left(\sum_{i=2}^{n} 1/i\right)-2(n-1)$。

f) 从 (e) 和 $\sum_{j=1}^{n} 1/j \approx \ln n + \gamma$ 的事实证明快速排序算法用的平均比较次数是 $\Theta(n\log n)$。这里的 $\gamma = 0.577\,21\cdots$ 是欧拉常数。

*35. 所谓固定元素就是排序后仍旧保持在原来位置的元素。在随机选择的 n 个元素的排列中，固定元素个数的方差是什么？[提示：设 X 表示一个随机排列中的固定元素个数。写出 $X=X_1+X_2+\cdots+X_n$，如果这个排列固定第 i 个元素，则 $X_i=1$，否则为 0。]

两个随机变量 X 和 Y 在样本空间 S 上的协方差记作 $\mathrm{Cov}(X, Y)$，定义为随机变量 $(X-E(X))(Y-E(Y))$ 的期望值，即 $\mathrm{Cov}(X, Y)=E((X-E(X))(Y-E(Y)))$。

36. 证明 $\mathrm{Cov}(X, Y)=E(XY)-E(X)E(Y)$，并使用这一结果证明，如果 X 和 Y 是独立的随机变量则 $\mathrm{Cov}(X, Y)=0$。

37. 证明 $V(X+Y)=V(X)+V(Y)+2\mathrm{Cov}(X, Y)$。

38. 如果 X 和 Y 是具有 $X((i, j))=2i$ 和 $Y((i, j))=i+j$ 的随机变量，其中 i 和 j 是掷两个均匀的骰子时出现在第一和第二个骰子上的点数，求 $\mathrm{Cov}(X+Y)$。

39. m 个球被均匀地随机分到 n 个箱子里，使得第一个箱子空的概率是多少？

40. m 个球被均匀地随机分到 n 个箱子里，预期落入第一个箱子的球数是多少？

41. m 个球被均匀地随机分到 n 个箱子里，预期空箱子数是多少？

关键术语和结果

术语

事件的概率：该事件成功的结果次数除以可能结果的总次数。

$P(E \mid F)$（给定条件 F 下 E 的条件概率）：$p(E \cap F)/p(F)$。

独立事件：使得 $p(E\cap F)=p(E)p(F)$ 成立的事件 E 和 F。

随机变量：一个函数，它对一个试验的每次结果赋一个实数值。

随机变量 X 的分布：对 $(r, p(X=r))$ 的集合，这里 $r\in X(S)$。

均匀分布：对一个有穷集元素的等概率赋值。

随机变量的期望值：一个随机变量的加权平均，用结果的概率加权的随机变量的值，即 $E(X)=\sum_{s\in S} p(s)X(s)$。

几何分布：一个随机变量 X 的分布，且对某个 p，使得对 $k=1, 2, \cdots$ 有 $p(X=k)=(1-p)^{k-1}p$。

独立随机变量：随机变量 X 和 Y 使得对于所有的实数 r_1 和 r_2，有 $p(X(s)=r_1$ 且 $Y(s)=r_2)=p(X(s)=r_1)p(Y(s)=r_2)$。

随机变量的方差：随机变量的值与它的期望值之差平方的加权平均，其中的权由结果的概率给定，即 $V(X)=\sum_{s\in S}(X(s)-E(X))^2 p(s)$。

伯努利试验：一个具有两种可能结果的试验。

概率算法：做一步或多步随机选择的算法。

概率方法：证明与集合中具有给定性质个体的有关结果的一种证明技巧，它通过对个体指派概率，然后证明有着这种性质的个体的概率是正数。

结果

当执行 n 次独立的伯努利试验时，k 次成功的概率等于 $C(n, k)p^k q^{n-k}$，其中 p 是成功的概率且 $q=1-p$ 是失败的概率。

$$E(X) = \sum_{r \in X(S)} p(X = r) r$$

期望的线性性质：如果 X_1，X_2，…，X_n 是随机变量，$E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$。

如果 X 和 Y 是独立的随机变量，则 $E(XY) = E(X)E(Y)$。

如果 X_1，X_2，…，X_n 是独立的随机变量，则 $V(X_1 + X_2 + \cdots + X_n) = V(X_1) + V(X_2) + \cdots + V(X_n)$。

切比雪夫不等式：$p(|X(s) - E(X)| \geq r) \leq V(X)/r^2$。其中 X 是具有概率函数 p 的随机变量，r 是一个正实数。

复习题

1. a)当所有的结果是等可能时，定义一个事件的概率。
   b)在买彩票时从前 50 个正整数中选择 6 个不同的中奖数，那么买一张彩票选对 6 个中奖整数的概率是多少？
2. a)一个有限样本空间对结果所赋的概率应该满足什么条件？
   b)如果头像出现的次数是非头像的 3 倍，那么赋给头像和非头像结果的概率应该是多少？
3. a)定义给定事件 F 下事件 E 的条件概率。
   b)假设 E 是掷骰子时出现偶数点的事件，F 是掷骰子时出现 1，2 或 3 点的事件，那么给定 E 下 F 的概率是什么？
4. a)什么时候两个事件 E 和 F 是独立的？
   b)假设 E 是掷一个均匀的骰子时出现偶数点的事件，F 是出现 5 点或 6 点的事件，那么 E 和 F 是否独立？
5. a)什么是随机变量？
   b)设 X 是随机变量，它对掷两个骰子的事件所赋的值是两个骰子上较大的点数。哪些是这个随机变量的赋值？
6. a)定义随机变量 X 的期望值。
   b)设 X 是随机变量，它对掷两个骰子的事件所赋的值是两个骰子上较大的点数。那么随机变量 X 的期望值是什么？
7. a)解释怎样把具有有限多个可能输入的算法在平均情形下的计算复杂度转变成期望值。
   b)如果要找的元素在表中的概率是 1/3，并且这个元素是表中 n 个元素之一的可能性是相等的，那么线性搜索算法在平均情形下的计算复杂度是什么？
8. a)伯努利试验的含义是什么？
   b)在 n 次独立的伯努利试验中 k 次成功的概率是多少？
   c)在 n 次独立的伯努利试验中成功次数的期望值是什么？
9. a)什么是随机变量的期望的线性性质？
   b)当一个帽子寄存人随机发回帽子时，怎样使用期望的线形性质帮助我们找到能够收到自己帽子的预期人数？
10. a)如果能够接受一个小概率的错误时，怎样使用概率求解一个判定问题？
    b)如果愿意接受一个小概率的出错时，我们怎样快速确定一个正整数是否为素数？
11. a)说一个随机变量具有概率 p 的几何分布是什么意思？
    b)具有概率 p 的几何分布的平均值是什么？
12. a)什么是随机变量的方差？
    b)具有成功概率为 p 的伯努利试验的方差是什么？

13. a)什么是 n 个独立随机变量的和的方差？

b)设每次试验的成功概率为 p，当执行 n 次独立的伯努利试验时，成功次数的方差是什么？

14. 对于一个随机变量与它的平均值的偏差超过一个指定量的概率，切比雪夫不等式告诉我们什么结果？

补充练习

1. 从 1 到 40 之间(含 1 和 40 在内)选出 6 个连续的数作为彩票中奖号码的概率是多少？
2. 一手 13 张牌不包含对的概率是多少？
3. 求下述各种情况下的概率。选一手 13 张桥牌包含

a)全部 13 张红心。

b)同种花色的 13 张牌。

c)7 张黑桃和 6 张梅花。

d)一种花色的 7 张牌和另一种花色的 6 张牌。

e)4 张方块、6 张红心、2 张黑桃和 1 张梅花。

f)一种花色的 4 张牌，第二种花色的 6 张牌，第三种花色的 2 张牌，第四种花色的 1 张牌。

4. 求下述各种情况下的概率。选 7 张扑克牌包含

a)2 类，其中一类 4 张，第二类 3 张。

b)3 类，其中一类 3 张，另外两类每类各 2 张。

c)4 类，其中三类每类各 2 张，第四类 1 张。

d)5 类，其中两类每类 2 张，另三类每类 3 张。

e)7 类不同的牌。

f)一个 7 张牌的同花。

g)一个 7 张牌的顺子。

h)一个 7 张牌的同花顺子。

一个八面体骰子有 8 个面，面上的数字为 1 到 8。

5. a)当掷一个均匀的八面体骰子时，出现数字的期望值是什么？

b)当掷一个均匀的八面体骰子时，出现数字的方差是什么？

一个十二面体骰子有 12 个面，面上的数字为 1 到 12。

6. a)当掷一个均匀的十二面体骰子时，出现数字的期望值是什么？

b)当掷一个均匀的十二面体骰子时，出现数字的方差是什么？

7. 假设掷一对均匀的八面体骰子。

a)出现数字之和的期望值是什么？

b)出现数字之和的方差是什么？

8. 假设掷一对均匀的十二面体骰子。

a)出现数字之和的期望值是什么？

b)出现数字之和的方差是什么？

9. 假设一个均匀的标准骰子(立方体)和一个均匀的八面体骰子一起掷。

a)出现数字之和的期望值是什么？

b)出现数字之和的方差是什么？

10. 假设一个均匀的八面体骰子和一个均匀的十二面体骰子一起掷。

a)出现数字之和的期望值是什么？

b)出现数字之和的方差是什么？

11. 假设 n 个人($n\geqslant 3$)玩“单人出局”的游戏确定下一次谁买饮料，n 个人同时掷均匀的硬币，每人一个。如果除了一个以外其余所有硬币的结果都相同，那么这个掷出不同结果的人将买饮料。否则，这些人再次掷硬币，直到出现一个硬币与其他所有的硬币结果不同为止。

a)仅仅掷一次硬币，这个人就能确定下来的概率是多少？

b)第 k 次掷硬币时，这个人确定下来的概率是多少？

c)从 n 个人中确定这个人预期需要掷多少次硬币？

12. 设 p 和 q 是素数且 $n=pq$。随机选择小于 n 的正整数不被 p 或 q 整除的概率是多少？

*13. 设 m 和 n 是正整数。随机选择小于 mn 的正整数不被 m 或 n 整除的概率是多少？

14. 设 E_1，E_2，…，E_n 是 n 个事件满足 $p(E_i)>0$，$i=1, 2, \cdots, n$。证明

$$p(E_1 \cap E_2 \cap \cdots \cap E_n) = p(E_1)p(E_2 \mid E_1)p(E_3 \mid E_1 \cap E_2)\cdots p(E_n \mid E_1 \cap E_2 \cap \cdots \cap E_{n-1})$$

15. 在盒子里有 3 张卡片，一张卡片的两面是黑色的，一张卡片的两面是红色的，第三张卡片的一面是黑

色的，一面是红色的。我们随机取一张卡片并且只看它的一个面。

a)如果这个面是黑色的，另一个面也是黑色的概率是多少?

b)另一个面与我们看到的面是同色的概率是多少?

16. 当一个均匀的硬币被掷 n 次时，头像向上和头像向下出现的次数相等的概率是多少?

17. 一个随机选择的 10 位二进制串出现回文的概率是多少?

18. 一个随机选择的 11 位二进制串出现回文的概率是多少?

19. 考虑下面的游戏。一个人重复掷一个硬币直到头像在上。如果在掷第 n 次时第一次出现头像在上，他就得到 2^n 美元。

a)设 X 是等于这个人所赢得钱数的随机变量。证明 X 的期望值不存在(即它是无限的)。证明一个人应该心甘情愿地赌任意多的钱来玩这个游戏(这就叫做圣・匹茨堡悖论)。你觉得为什么它叫做悖论?

b)假设如果掷第 8 次之前出现第一个头像在上这个人就接受 2^n 美元的回报，如果在掷第 8 次或者第 8 次之后出现头像则接受 2^8 美元。这个人所赢钱数的期望值是什么? 为了玩这个游戏，一个人应该心甘情愿付多少钱?

20. 设将 n 个球抛进 b 个箱子使得每个球等可能地落入任何箱子并且这些抛放是独立的。

a)找出一个特定的球落入一个特定箱子的概率。

b)落入一个特定箱子的球的预期数目是多少?

c)直到一个特定的箱子包含一个球为止，预期需要投多少个球?

*d)直到所有的箱子都包含一个球为止，预期需要抛多少个球? [提示：设 X_i 表示一旦 $i-1$ 个箱子包含一个球，第 i 个箱子落入一个球所需要抛掷的球数。求 $E(X_i)$并且使用期望的线性性质。]

21. 设 A 和 B 是具有概率 $p(A)=3/4$，$p(B)=1/3$ 的事件。

a)最大的 $p(A\cap B)$可能是多少? 它最小可能是多少? 举例说明这两种极端情况对于 $p(A\cap B)$都是可能的。

b)最大的 $p(A\cup B)$可能是多少? 它最小可能是多少? 举例说明这两种极端情况对于 $p(A\cup B)$都是可能的。

22. 设 A 和 B 是具有概率 $p(A)=2/3$，$p(B)=1/2$ 的事件。

a)最大的 $p(A\cap B)$可能是多少? 它最小可能是多少? 举例说明这两种极端情况对于 $p(A\cap B)$都是可能的。

b)最大的 $p(A\cup B)$可能是多少? 它最小可能是多少? 举例说明这两种极端情况对于 $p(A\cup B)$都是可能的。

23. 我们说事件 E_1，E_2，…，E_n 是相互独立的，如果

$$p(E_{i_1}\cap E_{i_2}\cap\cdots\cap E_{i_m})=p(E_{i_1})p(E_{i_2})\cdots p(E_{i_m})$$

其中 i_j 是整数，$j=1,2,\cdots,m$，满足 $1\leqslant i_1<i_2<\cdots<i_m\leqslant n$ 且 $m\geqslant 2$。

a)写出三个事件 E_1，E_2，E_3 相互独立所要求的条件。

b)令 E_1，E_2 和 E_3 分别是掷硬币第 1 次出现头像、第 2 次不出现头像和第 3 次不出现头像的事件。当一个均匀的硬币被掷 3 次时，E_1，E_2 和 E_3 相互独立吗?

c)令 E_1，E_2 和 E_3 分别是掷硬币第 1 次出现头像、第 3 次出现头像和出现偶数个头像的事件。当一个均匀的硬币被掷 3 次时，E_1，E_2 和 E_3 相互独立吗?

d)为证明 n 个事件是相互独立的，必须检查多少个条件?

*24. 设 E 和 F 是满足 $p(F)\neq 0$ 的事件。证明 E 的概率是给定 F 下 E 的概率与给定 F 的补 $\overline{F}$ 下 E 的概率的加权平均，其中的权分别是 F 和 $\overline{F}$ 的概率，即

$$p(E)=p(E\mid F)p(F)+p(E\mid\overline{F})p(\overline{F})$$

[提示：利用 $E=(E\cap F)\cup(E\cap\overline{F})$]

*25. 设 E 是样本空间 S 的一个事件且 F_1，F_2，…，F_n 是使得 $\bigcup_{i=1}^{n}F_i=S$ 的相互排斥的事件。假定 $p(E)\neq 0$ 且

$p(F_i)\neq 0$，$i=1$，2，…，n。证明：

$$p(F_j \mid E)=\frac{p(E \mid F_j)p(F_j)}{\sum_{i=1}^{n} p(E \mid F_i)p(F_i)}$$

[提示：使用事实 $E=\bigcup_{i=1}^{n}(E\cap F_i)$。] 这一结果叫做贝叶斯公式，因为它是由英国哲学家托马斯·贝叶斯发现的。

*26. 海王星附近的一个空间探测器使用二进制串与地球通信。假设它发送 1 个 1 用时 1/3，发送 1 个 0 用时 2/3。当发送 0 时，被正确接收的概率是 0.9，而不被正确接收(接收成 1)的概率是 0.1。发送 1 时被正确接收的概率是 0.8，而不被正确接收(接收成 0)的概率是 0.2。

a)用练习 24 找到 1 个 0 被接收的概率。

b)用练习 25 给出的贝叶斯公式找到给定 1 个 0 被接收的条件下，0 被传送的概率。

27. 设 X 是样本空间 S 上的随机变量。证明 $V(aX+b)=a^2V(X)$，其中 a 和 b 是实数。

28. 用切比雪夫不等式证明不管多少个人寄存他们的帽子，当帽子寄存人随机返回帽子时有多于 10 个人取回自己帽子的概率不超过 1/100。(提示：使用例 6 和 5.3 节练习 35。)

29. 假设至少一个事件 $E_j(j=1, 2, \cdots, m)$保证出现，且不可能有多于两个事件出现。证明如果对 $j=1$，2，…，m 有 $p(E_j)=q$ 且对 $1\leqslant j<k\leqslant m$ 有 $p(E_j\cap E_k)=r$，那么 $q\geqslant 1/m$ 且 $r\leqslant 2/m$。

30. 证明如果 m 是正整数，那么当执行每次成功概率为 p 的独立的伯努利试验时，在第 $m+n$ 次试验出现第 m 次成功的概率是 $\binom{n+m-1}{n}q^n p^m$。

31. 当买某种产品时你可以得到 n 种不同的收藏卡作为奖品。假设每次你买这种产品时得到任何类型的收藏卡是等可能的。设 X 是随机变量，它等于每种类型的收藏卡至少得到一张而需购买的产品数，X_j 是随机变量，它等于已经收集到 j 种不同的卡以后直到再得到一张新卡所必须再购买的产品数，$j=0$，1，…，$n-1$。

a)证明 $X=\sum_{j=0}^{n-1}X_j$。

b)证明在 j 种不同的卡已经得到以后，再一次买产品时，得到一张新类型的卡具有的概率是$(n-j)/n$。

c) 证明 X_j 具有参数$(n-j)/n$ 的几何分布。

d) 使用 a) 和 c) 证明 $E(X)=n\sum_{j=1}^{n}1/j$。

e) 如果有 50 种不同类型的卡，使用近似式$\sum_{j=1}^{n}1/j\approx \ln n+\gamma$(其中 $\gamma=0.577\,21\cdots$，是欧拉常数)。求为得到每种类型的卡，你需要购买产品的预期数量。

32. 极大可满足性问题　一个复合命题是以合取范式给出(即这个命题表示为子句的合取，每个子句是两个或更多的变量或者它们的非的析取式)。求这个复合命题中一组真值变量的赋值以使得尽可能多的子句为真。例如在

$$(p\vee q)\wedge(p\vee\neg q)\wedge(\neg p\vee r)\wedge(\neg p\vee\neg r)$$

的子句中，通过对 p，q，r 的真值赋值可以有三个为真，但是没有四个为真。我们将证明用概率方法可以提供由于变量的真值赋值而可能为真的子句个数的一个下界。

a)假设在一个合取范式的复合命题中有 n 个变量。如果我们通过掷硬币随机对每个变量赋值，若硬币出现头像在上对变量赋真值，若头像在下则赋假值。对 n 个变量而言，每种可能赋值的概率是什么？

b)假设每个子句是恰好两个不同变量或者它们的非的析取式，给定 a)中随机的真值赋值，一个给定子句为真的概率是多少？

c)假设在复合命题中有 D 个子句。给这些变量随机地进行真值赋值，这些子句中预期为真的个数是

多少？

d)用(c)证明对每个合取范式的复合命题，存在对变量的一组真值赋值，使得至少 3/4 的子句为真。

计算机题目

按下述给定的输入和输出写程序。

1. 给定实数 p，$0\leqslant p\leqslant 1$，生成来自具有概率 p 的伯努利分布的随机数。
2. 给定正整数 n，生成集合$\{1, 2, 3, \cdots, n\}$的随机排列(参见 4.6 节练习 16)。
3. 给定正整数 m 和 n，生成前 n 个正整数的 m 个随机排列。求出在每个排列中的逆序数并确定这些逆序的平均个数。
4. 一个不均匀的硬币其头像在上出现的概率为 p。将这个硬币被重复掷 n(n 为正整数)次，并确定出现头像在上的次数。显示累积的结果。
5. 给定正整数 n 和 m，生成前 n 个正整数的 m 个随机排列。使用插入排序排序每一个排列，并计数用到的比较次数。确定在所有 m 个排列所使用的平均比较次数。
6. 给定正整数 n 和 m，生成前 n 个正整数的 m 个随机排列。使用冒泡排序的下述版本排序每个排列，当一次巡回完成而没有交换时结束，计数用过的比较次数。确定在所有 m 个排列使用的平均比较次数。
7. 给定正整数 m，模拟由于买产品而收集得到的卡片，以找出为得到全套 m 种不同的收藏卡必须买的产品数量(见补充练习 31)。
8. 给定正整数 m 和 n，模拟 n 个关键字的放置，其中具有关键字 k 的一个记录放在地址 $h(k)=k \bmod m$，并且确定是否至少存在一次碰撞。
9. 给定正整数 n，找出从集合$\{1, 2, \cdots, n\}$选中的 6 个数就是被机器选出的中彩号码的概率。
10. 模拟蒙地厅大厦 3 门问题(参见 5.1 节例 11)的重复试验来计算每种策略的胜率。

计算和研究

使用一个计算程序或你已经编写的程序做下面的练习。

1. 找出一手 5 张扑克牌的各种类型的概率并且根据它们的概率排列这些类型。
2. 找出在新泽西六合彩中买 1 美元奖票有大于 1 美元的中奖期望值的条件。为了赢奖，不管数的次序，你必须从 1 到 48 的正整数中(含 1 和 48 在内)选中被抽出的 6 个数。奖金在中奖的人中是平均分配的。必须考虑抽奖的奖金总额和买奖票的人数。
3. 由测试大量随机选择的整数对来估计随机选择的 2 个整数是互素的这一事件的概率。查找给出这个概率的定理并将你的结果与正确的概率作比较。
4. 确定需要多少人，才能保证其中至少 2 个人的生日在每年的同一天的概率至少是 70%，80%，90%，95%，98%和 99%。
5. 生成前 100 个正整数集合的 100 个随机选择的排列表(参见 4.6 节练习 16)。
6. 模拟 n 个人的单人出局过程(在补充练习的练习 11 中描述的)，其中 $3\leqslant n\leqslant 10$。对每个 n 值做大量试验并且用这些结果估计为找到这个出局的单人预期需要掷硬币的次数。你的结果与 5.2 节练习 29 的结果相符吗？假设恰好一个人有一个不均匀的硬币，其头像出现概率 $p\neq 0.5$，那么结果又是怎样呢？
7. 给定正整数 n，模拟一个帽子寄存员随机把帽子发给寄存人的过程。确定得到自己帽子的人数。

写作题目

用本书以外的资料，按下列要求写成短文。

1. 描述概率论的起源和它的早期应用。
2. 描述玩轮盘赌时你可能下的不同的赌注。找出这些赌注在美国玩法(即轮盘包含数 0 和 00 在内)的概率。对你来说什么是最好的赌注？什么是最坏的赌注？

3. 讨论当你玩 21 点的纸牌游戏和卡西诺纸牌游戏时赢的概率。对于在赌场下注的人是否存在一种赢的策略？
4. 调查双骰子赌博游戏并且讨论投手赢的概率，这个游戏和一个公平的游戏有多接近？
5. 讨论著名的牛顿-佩皮斯问题的历史和解答，这个问题问下述情况中哪种最可能出现：当 6 个骰子被掷时至少出现一个 6 点，当 12 个骰子被掷时至少出现 2 个 6 点，或者当 18 个骰子被掷时至少出现 3 个 6 点？
6. 解释厄多斯和任伊是怎样首先使用概率方法的，并且描述这个方法的一些其他应用。
7. 讨论概率算法的不同类型并且描述每种类型的某些实例。



第6章 高级计数技术

许多计数问题用第 4 章讨论的方法是不容易求解的。一个这样的问题就是：有多少个 n 位二进位串不包含两个连续的 0？为求解这个问题，令 a_n 是这种 n 位二进位串数。可以证明 $a_{n+1}=a_n+a_{n-1}$。这个等式叫做递推关系，它和初始条件 $a_1=2$ 和 $a_2=3$ 确定了序列 $\{a_n\}$。此外，从这个与序列的项有关的等式可以找到 a_n 的显式公式。正如我们将要看到的，可以用一种类似的技术来求解许多不同的计数问题。

我们也将看到，可以用形式幂级数(也叫做生成函数)来求解许多计数问题，其中 x 的幂的系数代表我们感兴趣的序列的项。除了求解计数问题，生成函数还可用于求解递推关系以及证明组合恒等式。

许多其他类型的计数问题不能使用第 4 章所讨论的技术求解，例如：有多少种方式把 7 项工作分给 3 个雇员而使得每个雇员至少得到一项工作？有多少个素数小于 1000？可以用计数集合并集中的元素数来求解这两个问题。我们将开发一种叫做容斥原理的技术来计数在集合并集中的元素个数，并且将说明怎样用这种技术求解计数问题。

可以用本章学到的技术与第 4 章的基本技术一起求解许多计数问题。

6.1 递推关系

6.1.1 引言

一群细菌的数目每小时增加一倍。如果开始有 5 个细菌，在 n 小时末将有多少个细菌？为求解这个问题，令 a_n 是 n 小时末的细菌数。因为细菌数每小时增加一倍，只要 n 是正整数，关系 $a_n=2a_{n-1}$ 就成立。对所有的非负整数 n，这个关系和初始条件 $a_0=5$ 一起唯一地确定了 a_n。利用这一信息可找出关于 a_n 的公式。

某些计数问题不能用第 4 章给出的技术求解，但可以通过找到序列的项之间的关系，如在涉及细菌的问题中的关系，即递推关系来求解。我们将研究各种能用递推关系构造模型的计数问题。我们也将在本节和下一节建立一些方法，针对满足某类递推关系的序列，求出序列的项的显式公式。

6.1.2 递推关系

第 3 章介绍了怎样递归定义一个序列。一个序列的递归定义指定了一个或多个初始的项以及一个由前项确定后项的规则。可以用递归定义来求解计数问题。对于递归定义，这个从某些前项求后项的规则就叫做递推关系。

定义 1 关于序列 $\{a_n\}$ 的递推关系是一个等式，它把 a_n 用序列中在 a_n 前面的一项或多项即 a_0，a_1，…，a_{n-1} 来表示，这里 $n\geqslant n_0$，n_0 是一个非负整数。如果一个序列的项满足递推关系，这个序列就叫做递推关系的解。

在递归和递推关系之间存在重要的联系，关于这一点我们在本章的后边再做研究。递归算法依据较小规模的同一问题的一个或者多个实例的解得到规模为 n 的问题的解。因此，当我们分析一个递归算法的复杂性时，我们得到一个递推关系，它把求解规模为 n 的问题需要的运算次数用求解一个或多个小规模实例的同一问题所需要的运算次数来表示。

例 1 令$\{a_n\}$是一个序列，它满足递推关系 $a_n=a_{n-1}-a_{n-2}$，$n=2,3,4,\cdots$，且 $a_0=3$，$a_1=5$，那么 a_2 和 a_3 是什么？

解 从递推关系可以看出，$a_2=a_1-a_0=5-3=2$ 且 $a_3=a_2-a_1=2-5=-3$。■

例 2 确定序列$\{a_n\}$是否为递推关系 $a_n=2a_{n-1}-a_{n-2}$，$n=2,3,4,\cdots$，的解。这里 $a_n=3n$，n 是非负整数。对 $a_n=2^n$ 和 $a_n=5$ 也回答同一个问题。

解 假设对每一个非负整数 n，$a_n=3n$，那么对于 $n\geqslant 2$，可以看出 $2a_{n-1}-a_{n-2}=2[3(n-1)]-3(n-2)=3n=a_n$。于是，$\{a_n\}$是该递推关系的解，其中 $a_n=3n$。

假设对每一个非负整数 n，$a_n=2^n$。注意到 $a_0=1$，$a_1=2$，$a_2=4$。因为 $2a_1-a_0=2\cdot 2-1=3\neq a_2$，不难看出序列$\{a_n\}$不是该递推关系的解，其中 $a_n=2^n$。

假设对每一个非负整数 n，$a_n=5$，那么对于 $n\geqslant 2$ 有 $a_n=2a_{n-1}-a_{n-2}=2\cdot 5-5=5=a_n$。因此，$\{a_n\}$，$a_n=5$ 是该递推关系的解。■

序列的初始条件说明了在递推关系起作用的首项之前的那些项。例如，例 1 中的 $a_0=3$ 和 $a_1=5$ 是初始条件。递推关系和初始条件唯一地确定了一个序列。这是由于一个递推关系和初始条件一起提供了这个序列的递归定义。只要使用足够多次，序列的任何一项都可以从初始条件开始通过递推关系求出。但是对于某些特定类型的序列，可以有更好的方法通过它的递推关系和初始条件来计算它的项。我们将在本节和 6.2 节讨论这些方法。

6.1.3 用递推关系构造模型

我们可以使用递推关系构造各种问题的模型，例如找复利，计数岛上的兔子，确定汉诺塔难题的移动次数，以及计数具有确定性质的二进位串。

例 3 **复利** 假设一个人在银行的储蓄账户上存了 10 000 美元，年复利是 11%。那么在 30 年后账上将有多少钱？

Extra Examples

解 为求解这个问题，令 P_n 表示 n 年后账上的钱数。因为 n 年后账上的钱等于在 $n-1$ 年后账上的钱加上第 n 年的利息，易见序列$\{P_n\}$满足递推关系

$$P_n=P_{n-1}+0.11P_{n-1}=(1.11)P_{n-1}$$

初始条件是 $P_0=10\,000$。

我们可以使用迭代法找到关于 P_n 的公式。注意

$$\begin{aligned}P_1&=(1.11)P_0\\P_2&=(1.11)P_1=(1.11)^2P_0\\P_3&=(1.11)P_2=(1.11)^3P_0\\&\vdots\\P_n&=(1.11)P_{n-1}=(1.11)^nP_0\end{aligned}$$

当代入初始条件 $P_0=10\,000$，就得到公式 $P_n=(1.11)^n 10\,000$。我们可以使用数学归纳法验证它的正确性。公式对 $n=0$ 是正确的，这是初始条件的直接结果。假定 $P_n=(1.11)^n 10\,000$，那么由递推关系和归纳假设，

$$P_{n+1}=(1.11)P_n=(1.11)(1.11)^n 10\,000=(1.11)^{n+1}10\,000$$

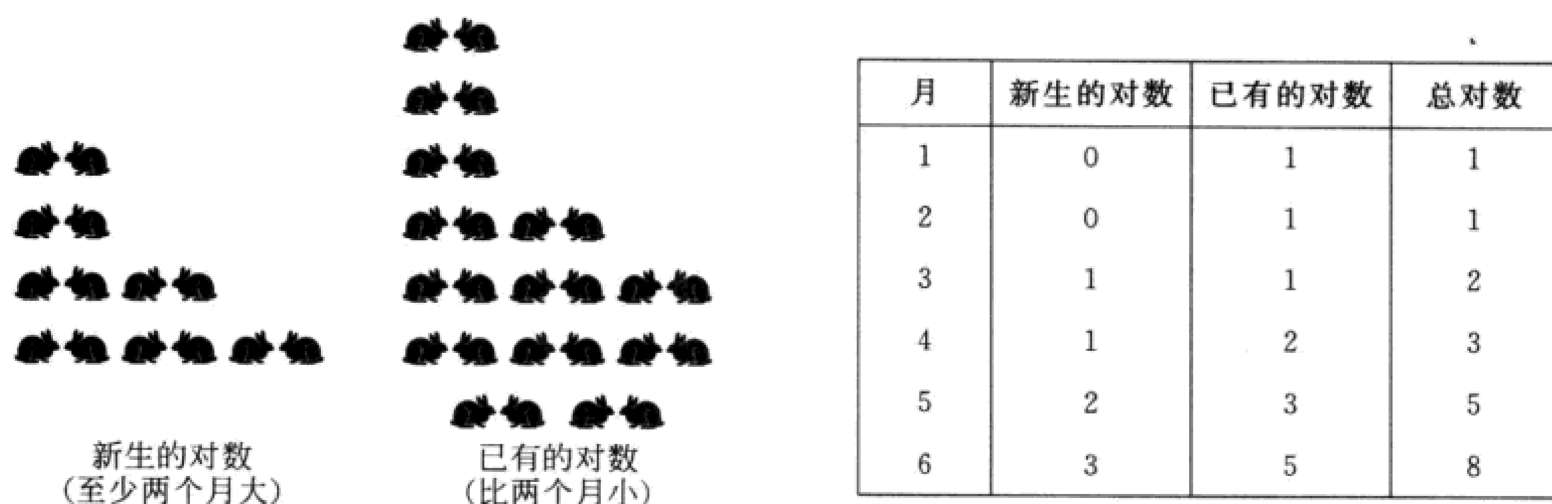
这证明了对 P_n 的显式公式是正确的。

将 $n=30$ 代入公式 $P_n=(1.11)^n 10\,000$ 就证明了在 30 年后账上包含

$$P_{30}=(1.11)^{30}10\,000=228\,922.97 \text{ 美元}$$ ■

例 4 说明了怎样用递推关系建立关于岛上兔子数的模型。

例 4　兔子和斐波那契数　考虑下面的问题，它是由里奥那多·底·比萨，也就是斐波那契，于 13 世纪在《算书》(Liber abaci)一书中提出来的。一对刚出生的兔子(一公一母)被放到岛上。每对兔子出生后两个月才开始繁殖后代。如图 6-1 所示，在出生两个月以后，每对兔子在每个月都将繁殖一对新的兔子。假定兔子不会死去，找出 n 个月后关于岛上兔子对数的递推关系。



月	新生的对数	已有的对数	总对数
1	0	1	1
2	0	1	1
3	1	1	2
4	1	2	3
5	2	3	5
6	3	5	8

图 6-1　岛上的兔子

解　用 f_n 表示 n 个月后的兔子对数。我们将证明 f_n，$n=1, 2, 3, \cdots$是斐波那契序列的项。

可以用递推关系建立兔子数的模型。在第 1 个月末，岛上的兔子对数是 $f_1=1$。由于这对兔子在第 2 个月没有繁殖，因此 $f_2=1$。为找到 n 个月后的兔子对数，要把前一个月岛上的对数 f_{n-1}加上新生的对数，而这个数等于 f_{n-2}，因为每对两个月大的兔子都生出一对新兔子。

因此，序列$\{f_n\}$满足递推关系

$$f_n = f_{n-1} + f_{n-2}, \quad n \geqslant 3$$

和初始条件 $f_1=1$ 和 $f_2=1$。由于这个递推关系和初始条件唯一地确定了这个序列，因此 n 个月后岛上的兔子对数由第 n 个斐波那契数给出。 ■

例 5 涉及一个著名的难题。

例 5　汉诺塔　19 世纪后期由法国数学家埃德沃德·卢卡斯发明的一个流行的游戏叫做汉诺塔，它由安装在一个板上的 3 根柱子和若干大小不同的盘子构成。开始时，这些盘子按照大小的次序放在第一根柱子上，使得大盘子在底下(如图 6-2 所示)。游戏的规则是：每一次把 1 个盘子从一根柱子移动到另一根柱子，但是不允许这个盘子放在比它小的盘子上面。游戏的目标是把所有的盘子按照大小的次序都放到第二根柱子上，并且将最大的盘子放在底部。

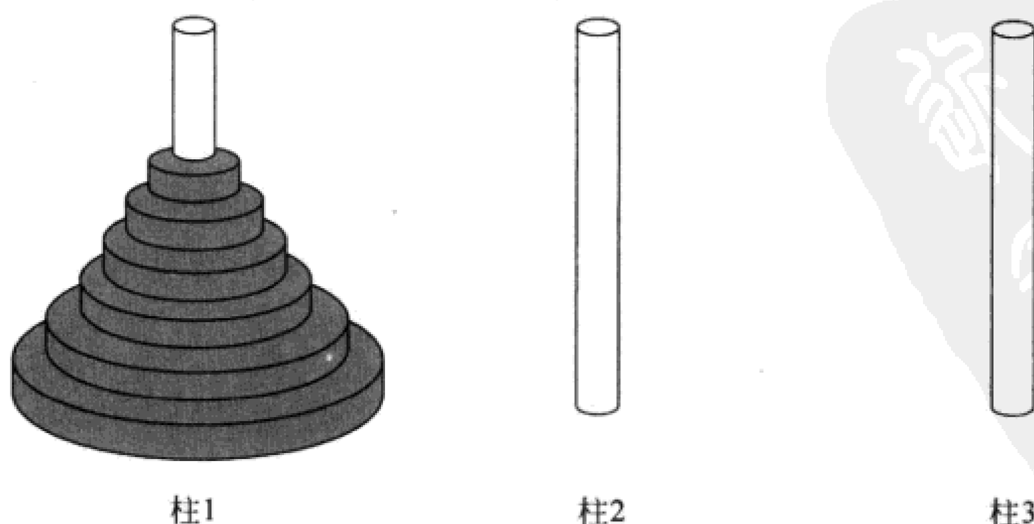


图 6-2　汉诺塔的初始位置

令 H_n 表示解 n 个盘子的汉诺塔问题所需要的移动次数。建立一个关于序列 $\{H_n\}$ 的递推关系。

解 开始 n 个盘子在柱 1。按照游戏规则我们可以用 H_{n-1} 次移动将上边的 $n-1$ 个盘子移到柱 3(图 6-3 说明了此刻的柱子和盘子)。在这些移动中保留最大的盘子不动。然后，我们用一次移动将最大的盘子移到第二根柱子上。我们可以再使用 H_{n-1} 次移动将柱 3 上的 $n-1$ 个盘子移到柱 2，把它们放到最大的盘子上面，这个最大的盘子一直放在柱 2 的底部。容易看出，使用更少的步数是不可能求解这个难题的。这就证明了

$$H_n = 2H_{n-1} + 1$$

初始条件是 $H_1 = 1$，因为依照规则一个盘子可以用 1 次移动从柱 1 移到柱 2。

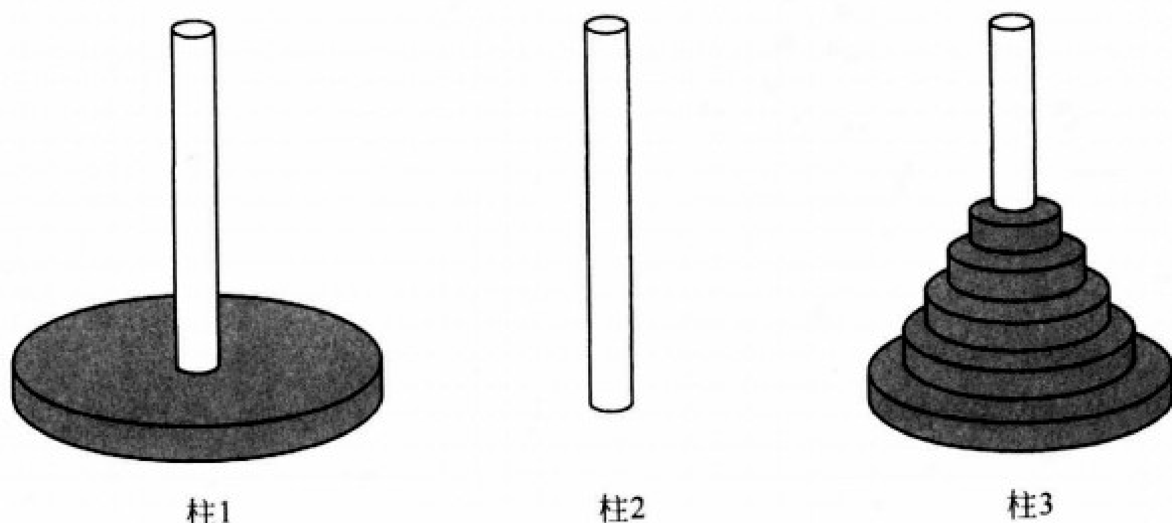


图 6-3 汉诺塔的一个中间位置

我们可以使用迭代方法求解这个递推关系。

$$\begin{aligned} H_n &= 2H_{n-1} + 1 \\ &= 2(2H_{n-2} + 1) + 1 = 2^2 H_{n-2} + 2 + 1 \\ &= 2^2(2H_{n-3} + 1) + 2 + 1 = 2^3 H_{n-3} + 2^2 + 2 + 1 \\ &\vdots \\ &= 2^{n-1} H_1 + 2^{n-2} + 2^{n-3} + \cdots + 2 + 1 \\ &= 2^{n-1} + 2^{n-2} + \cdots + 2 + 1 \\ &= 2^n - 1 \end{aligned}$$

为了用序列前面的项表示 H_n，我们重复地用到这个递推关系。在倒数第二个等式中用了初始条件 $H_1 = 1$。最后一个等式是基于等比序列的求和公式得出的，这个公式可以在 3.2 节的定理 1 中找到。

用迭代方法找出了具有初始条件 $H_1 = 1$ 的递推关系 $H_n = 2H_{n-1} + 1$ 的解。这个公式可以用数学归纳法证明。证明留给读者作为节后的练习。

关于这个难题有一个古老的传说，在汉诺有一座塔，那里的僧侣按照这个游戏的规则从一个柱子到另一个柱子移动 64 个金盘子，据说当他们结束游戏时世界就到了末日。如果这些僧侣 1 秒钟移动 1 个盘子，这个世界将在他们开始多久以后终结？

根据这个显式公式，僧侣需要

$$2^{64} - 1 = 18\ 446\ 744\ 073\ 709\ 551\ 615$$

次移动来搬这些盘子。每次移动需要 1 秒钟，他们将用 5000 亿年来求解这个难题，因此这个世界的寿命应该相当长。 ■

注意 许多人研究了源自例 5 所述汉诺塔难题的各种问题。某些问题用到更多的柱子，某些问题允许同样大小的盘子，某些问题对盘子的移动类型加以限制。一个最古老和最有趣的问题是**雷夫难题**⊖，它是 1907 年由亨利·杜德尼在他的《坎特伯雷迷题》(*Canterbury Puzzle*)一书中提出来的。这个难题是雷夫出的，他让一个朝圣者把一堆各种大小的奶酪从 4 个凳子中的一个移到另一个，移动中不允许把直径较大的奶酪放在较小的奶酪上面。如果用柱子和盘子的概念来表述雷夫难题，除了使用 4 根柱子之外，其他和汉诺塔的规则一样。你可能会奇怪没有人能够确定求解 n 个盘子的雷夫难题所需要的最少移动次数。但是，存在一个猜想，至今已经超过 50 年了。这个猜想认为所需移动的最少次数等于由弗雷姆(Frame)和斯图尔特(Stewart)在 1939 年发明的算法所使用的移动次数。(更详细的信息参见节末的练习 54～61 和[St94])

例 6 说明了怎样用递推关系计数具有指定长度和某种性质的二进位串。

例 6 对于不含 2 个连续 0 的 n 位二进位串的个数找出递推关系和初始条件。有多少个这样的 5 位二进位串？

解 设 a_n 表示不含 2 个连续 0 的 n 位二进位串数。为得到一个关于 $\{a_n\}$ 的递推关系，由求和法则，不含 2 个连续 0 的 n 位二进位串的个数等于以 0 结尾的这种二进位串数加上以 1 结尾的这种二进位串数。我们将假定 $n \geqslant 3$，使得二进位串至少有 3 位。

精确地说，不含 2 个连续 0 并以 1 结尾的 n 位二进位串就是在不含 2 个连续 0 的 $n-1$ 位二进位串的尾部加上一个 1。因此存在 a_{n-1} 个这样的二进位串。

不含 2 个连续 0 并以 0 结尾的 n 位二进位串在它们的 $n-1$ 位必须是 1；否则它们就将以 2 个 0 结尾。因而，精确地说，不含 2 个连续 0 并以 0 结尾的 n 位二进位串就是在不含 2 个连续 0 的 $n-2$ 位二进位串的尾部加上 10。因此存在 a_{n-2} 个这样的二进位串。

如图 6-4 所示，可以断言对于 $n \geqslant 3$ 有

$$a_n = a_{n-1} + a_{n-2}$$

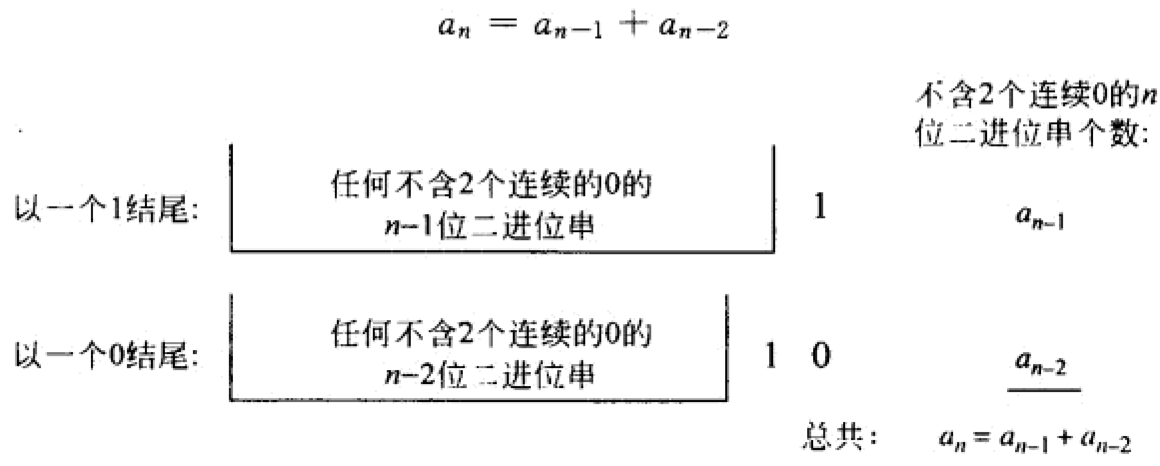


图 6-4 计数不含 2 个连续的 0 的 n 位二进位串

初始条件是 $a_1 = 2$，因为 1 位的二进位串是 0 和 1，没有连续的 2 个 0。而 $a_2 = 3$，因为 2 位的二进位串中满足条件的是 01，10 和 11。使用 3 次递推关系就可得到 a_5，

$$a_3 = a_2 + a_1 = 3 + 2 = 5$$
$$a_4 = a_3 + a_2 = 5 + 3 = 8$$
$$a_5 = a_4 + a_3 = 8 + 5 = 13$$

■

注意 $\{a_n\}$ 和斐波那契序列满足同样的递推关系。因为 $a_1 = f_3$ 且 $a_2 = f_4$，从而有

⊖ 雷夫，Reve，更常见的是拼写为 reeve，这个词在古代是指地方长官(governor)。

$a_n = f_{n+2}$。

例 7 说明怎样用递推关系建立编码字数的模型，这种编码字是通过确定的正确性检测所允许的。

例 7　编码字的枚举　一个计算机系统把一个十进制数字串作为一个编码字，如果它包含偶数个 0，就是有效的。例如，1 230 407 869 是有效的，而 120 987 045 608 不是有效的。设 a_n 是有效的 n 位编码字的个数。找出一个关于 a_n 的递推关系。

解　注意到 $a_1=9$，因为存在 10 个 1 位十进制数字串，并且只有一个，即串 0 是无效的。通过考虑怎样由 $n-1$ 位的数字串构成一个 n 位有效数字串，就可以推导出关于这个序列的递推关系。从少 1 位数字的串构成 n 位有效数字串有两种方式。

第一种，在一个 $n-1$ 位的有效数字串后面加上一个非 0 的数字就可以得到一个 n 位的有效数字串。加这个数字的方式有 9 种。因此用这种方法构成 n 位有效数字串的方式有 $9a_{n-1}$种。

第二种，在一个无效的 $n-1$ 位数字串后面加上一个 0 就可以得到 n 位有效的数字串。(这将产生具有偶数个 0 的串，因为无效的 $n-1$ 位数字串有奇数个 0。)这样做的方式数等于无效的 $n-1$ 位数字串的个数。因为存在 10^{n-1}个 $n-1$ 位数字串，其中有 a_{n-1}个是有效的，通过在无效的 $n-1$位数字串后面加上一个 0 就得到 $10^{n-1}-a_{n-1}$个 n 位的有效数字串。

因为所有的 n 位有效数字串都用这两种方式之一产生，从而存在

$$\begin{aligned} a_n &= 9a_{n-1} + (10^{n-1} - a_{n-1}) \\ &= 8a_{n-1} + 10^{n-1} \end{aligned}$$

个 n 位有效数字串。■

例 8 中的递推关系在许多不同的场合都可以见到。

例 8　求关于 C_n 的递推关系，其中 C_n 是通过对 $n+1$ 个数 x_0，x_1，x_2，…，x_n 的乘积中加括号来规定乘法的次序的方式数。例如，$C_3=5$，因为对 $x_0\cdot x_1\cdot x_2\cdot x_3$ 有 5 种加括号的方式来确定乘法的次序：

$$((x_0\cdot x_1)\cdot x_2)\cdot x_3 \quad (x_0\cdot(x_1\cdot x_2))\cdot x_3 \quad (x_0\cdot x_1)\cdot(x_2\cdot x_3)$$
$$x_0\cdot((x_1\cdot x_2)\cdot x_3) \quad x_0\cdot(x_1\cdot(x_2\cdot x_3))$$

解　为求得关于 C_n 的递推关系，我们注意到无论怎样在 $x_0\cdot x_1\cdot x_2\cdot\cdots\cdot x_n$ 中插入括号总有一个“·”运算符留在所有括号的外边，即执行最后一次乘法的运算符。[例如，在$(x_0\cdot(x_1\cdot x_2))\cdot x_3$ 中的最后一个运算符，在$(x_0\cdot x_1)\cdot(x_2\cdot x_3)$中的第二个运算符。]这个最后的运算符出现在 $n+1$ 个数中的两个数之间，比如说 x_k 和 x_{k+1}之间。当最后的运算符出现在 x_k 和 x_{k+1}之间时，存在 C_kC_{n-k-1}种方式插入括号来确定 $n+1$ 个数被乘的次序，因为有 C_k 种方式在乘积 $x_0\cdot x_1\cdot\cdots\cdot x_k$ 中插入括号来确定这 $k+1$ 个数的乘法次序，且有 C_{n-k-1}种方式在乘积 $x_{k+1}\cdot x_{k+2}\cdot\cdots\cdot x_n$ 中插入括号来确定这 $n-k$ 个数的乘法次序。由于这个最后的运算符可能出现在 $n+1$ 个数的任两个数之间，所以

$$\begin{aligned} C_n &= C_0C_{n-1} + C_1C_{n-2} + \cdots + C_{n-2}C_1 + C_{n-1}C_0 \\ &= \sum_{k=0}^{n-1} C_kC_{n-k-1} \end{aligned}$$

注意到初始条件是 $C_0=1$ 和 $C_1=1$。这个递推关系可以用生成函数的方法求解，这种方法将在 6.4 节讨论。可以证明 $C_n=C(2n, n)/(n+1)$。(见节末的练习 41。)■



序列$\{C_n\}$是卡塔兰(Catalan)数的序列。这个序列是除上例之外的许多不同计数问题的解(细节见[MiRo91]或[Ro84a]中有关卡塔兰数的一章)。

练习

1. 求给定序列的前 5 项，每个序列由下面的递推关系和初始条件定义。

a) $a_n=6a_{n-1}$，$a_0=2$
b) $a_n=a_{n-1}^2$，$a_1=2$
c) $a_n=a_{n-1}+3a_{n-2}$，$a_0=1$，$a_1=2$
d) $a_n=na_{n-1}+n^2a_{n-2}$，$a_0=1$，$a_1=1$
e) $a_n=a_{n-1}+a_{n-3}$，$a_0=1$，$a_1=2$，$a_2=0$

2. 求出由下列每个递推关系和初始条件所定义序列的前 6 项。

a) $a_n=-2a_{n-1}$，$a_0=-1$
b) $a_n=a_{n-1}-a_{n-2}$，$a_0=2$，$a_1=-1$
c) $a_n=3a_{n-1}^2$，$a_0=1$
d) $a_n=na_{n-1}+a_{n-2}^2$，$a_0=-1$，$a_1=0$
e) $a_n=a_{n-1}-a_{n-2}+a_{n-3}$，$a_0=1$，$a_1=1$，$a_2=2$

3. 令 $a_n=2^n+5\cdot 3^n$，$n=0, 1, 2, \cdots$。

a) 求 a_0，a_1，a_2，a_3 和 a_4。
b) 证明 $a_2=5a_1-6a_0$，$a_3=5a_2-6a_1$，$a_4=5a_3-6a_2$。
c) 证明对一切整数 n，$n\geqslant 2$，$a_n=5a_{n-1}-6a_{n-2}$。

4. 证明序列 $\{a_n\}$ 是递推关系 $a_n=-3a_{n-1}+4a_{n-2}$ 的解，如果

a) $a_n=0$
b) $a_n=1$
c) $a_n=(-4)^n$
d) $a_n=2(-4)^n+3$

5. 序列 $\{a_n\}$ 是递推关系 $a_n=8a_{n-1}-16a_{n-2}$ 的解吗？如果

a) $a_n=0$
b) $a_n=1$
c) $a_n=2^n$
d) $a_n=4^n$
e) $a_n=n4^n$
f) $a_n=2\cdot 4^n+3n\cdot 4^n$
g) $a_n=(-4)^n$
h) $a_n=n^2 4^n$

6. 对下面每一个序列求出满足这个序列的递推关系。(答案不是唯一的，因为任何序列满足的递推关系有无数多个。)

a) $a_n=3$
b) $a_n=2n$
c) $a_n=2n+3$
d) $a_n=5^n$
e) $a_n=n^2$
f) $a_n=n^2+n$
g) $a_n=n+(-1)^n$
h) $a_n=n!$

7. 证明序列 $\{a_n\}$ 是递推关系 $a_n=a_{n-1}+2a_{n-2}+2n-9$ 的解，如果

a) $a_n=-n+2$
b) $a_n=5(-1)^n-n+2$
c) $a_n=3(-1)^n+2^n-n+2$
d) $a_n=7\cdot 2^n-n+2$

8. 使用例 5 中用到的迭代方法求下述每个递推关系在给定初始条件下的解。

a) $a_n=-a_{n-1}$，$a_0=5$
b) $a_n=a_{n-1}+3$，$a_0=1$
c) $a_n=a_{n-1}-n$，$a_0=4$
d) $a_n=2a_{n-1}-3$，$a_0=-1$
e) $a_n=(n+1)a_{n-1}$，$a_0=2$
f) $a_n=2na_{n-1}$，$a_0=3$
g) $a_n=-a_{n-1}+n-1$，$a_0=7$

9. 用例 5 中用到的迭代方法求下面每个递推关系和初始条件的解。

a) $a_n=3a_{n-1}$，$a_0=2$
b) $a_n=a_{n-1}+2$，$a_0=3$
c) $a_n=a_{n-1}+n$，$a_0=1$
d) $a_n=a_{n-1}+2n+3$，$a_0=4$
e) $a_n=2a_{n-1}-1$，$a_0=1$
f) $a_n=3a_{n-1}+1$，$a_0=1$
g) $a_n=na_{n-1}$，$a_0=5$
h) $a_n=2na_{n-1}$，$a_0=1$

10. 一个人在账上存入 1000 美元，每年的复利是 9%。

a) 对于 n 年后账上的钱数建立一个递推关系。
b) 对于 n 年后账上的钱数求出一个显式公式。
c) 在 100 年以后账上将有多少钱？

11. 假设一群细菌的数目每小时增长为 3 倍。

a) 建立关于 n 小时后细菌数的递推关系。
b) 如果初始的群体有 100 个细菌，那么 10 小时后将有多少个细菌？

12. 假设世界人口在2002年是62亿，每年的增长率1.3%。
    a)对于2002年后n年的世界人口建立一个递推关系。
    b)求出2002年后n年的世界人口的显式公式。
    c)在2022年世界的人口将是多少?
13. 一个工厂逐月增长地定做体育赛车。在第1个月只做了1辆，在第2个月做了2辆，照此下去，到第n个月做了n辆。
    a)对这个工厂前n个月生产的赛车数构造一个递推关系。
    b)在第一年生产了多少辆赛车?
    c)求出这个工厂在前n个月生产赛车数的显式公式。
14. 一个雇员在1999年进入一个公司，初始的年薪是50 000美元。每年这个雇员的收入增加1000美元外加前一年工资的5%。
    a)对这个雇员在1999年后n年的工资建立一个递推关系。
    b)这个雇员在2007年的工资是多少?
    c)求出这个雇员在1999年后n年工资的显式公式。
15. 借5000美元，利率7%，如果每个月还款100美元，找出关于第k个月末欠款余额$B(k)$的递推关系。[提示：将$B(k)$用$B(k-1)$项表示，月息是$(0.07/12)B(k-1)$。]
16. a)借款利率r，每个月还款P，找出关于第k个月末欠款余额$B(k)$的递推关系。[提示：将$B(k)$用$B(k-1)$项表示，月息是$r/12$。]
    b)确定月还款P应该是多少才能使得这笔贷款在T个月后还完?
17. 用数学归纳法验证在例5导出的求解汉诺塔难题所需移动次数的公式。
18. a)找到一个关于n元素集合的排列数的递推关系。
    b)通过迭代用这个递推关系求n元素集合的排列数。
19. 一台出售邮票簿的售货机只接受1美元硬币、1美元纸币以及5美元纸币。
    a)找出与放n美元到这台售货机的方式数有关的递推关系，这里要考虑硬币和纸币放入的次序。
    b)初始条件是什么?
    c)一本邮票簿需10美元，有多少种方式付款?
20. 一个国家使用的硬币价值为1比索、2比索、5比索和10比索，纸币的价值为5比索、10比索、20比索、50比索和100比索。如果考虑付硬币和纸币的次序，求一个与付n比索账单的方式数有关的递推关系。
21. 如果考虑付硬币和纸币的次序，那么使用练习20描述的货币系统付17比索的账单有多少种方式?

*22. a)设n是正整数，求一个与下述正整数序列的个数有关的递推关系。这种序列要以1作为首项，以n作为末项并且是严格递增的。即序列a_1，a_2，…，a_k，其中$a_1=1$，$a_k=n$，且对$j=1, 2, \cdots, k-1$，$a_j<a_{j+1}$。
    b)初始条件是什么?
    c)当n是大于等于2的正整数时，有多少个(a)中所描述的序列?

23. a)求与包含2个连续0的n位二进位串的个数有关的递推关系。
    b)初始条件是什么?
    c)包含2个连续0的7位二进位串有多少个?
24. a)求与包含3个连续0的n位二进位串的个数有关的递推关系。
    b)初始条件是什么?
    c)包含3个连续0的7位二进位串有多少个?
25. a)求与不包含3个连续0的n位二进位串的个数有关的递推关系。
    b)初始条件是什么?
    c)不包含3个连续0的7位二进位串有多少个?

*26. a)求与包含01的n位二进位串的个数有关的递推关系。

b)初始条件是什么?

c)包含 01 的 7 位二进位串有多少个?

27. a)一个人爬阶梯，如果每次可以上 1 或 2 阶，求与爬 n 步阶梯的方式数有关的递推关系。

b)初始条件是什么?

c)这个人爬 8 步阶梯上飞机有多少种方式?

28. a)如果一个人爬阶梯每次可以上 1，2 或 3 阶，求与爬 n 步阶梯的方式数有关的递推关系。

b)初始条件是什么?

c)这个人爬 8 步阶梯上飞机有多少种方式?

一个只包含 0，1 和 2 的串叫做三进位串。

29. a)求与不包含 2 个连续 0 的 n 位三进位串的个数有关的递推关系。

b)初始条件是什么?

c)不包含 2 个连续 0 的 6 位三进位串有多少个?

30. a)求与包含 2 个连续 0 的 n 位三进位串的个数有关的递推关系。

b)初始条件是什么?

c)包含 2 个连续 0 的 6 位三进位串有多少个?

*31. a)求与不包含 2 个连续 0 或 2 个连续 1 的 n 位三进位串的个数有关的递推关系。

b)初始条件是什么?

c)不包含 2 个连续 0 或 2 个连续 1 的 6 位三进位串有多少个?

*32. a)求与包含 2 个连续 0 或 2 个连续 1 的 n 位三进位串的个数有关的递推关系。

b)初始条件是什么?

c)包含 2 个连续 0 或 2 个连续 1 的 6 位三进位串有多少个?

*33. a)求与不包含连续的相同符号的 n 位三进位串的个数有关的递推关系。

b)初始条件是什么?

c)不包含连续的相同符号的 6 位三进位串有多少个?

**34. a)求包含 2 个连续的相同符号的 n 位三进位串的个数的递推关系。

b)初始条件是什么?

c)包含 2 个连续的相同符号的 6 位三进位串有多少个?

35. 信息通过信道传送要使用两个信号。一个信号的传送需要 1 微秒，而另一个信号的传送需要 2 微秒。

a)求与在 n 微秒内发送的不同信息数有关的递推关系，其中信息由这两个信号的序列构成，并且信息中的每个信号后面都紧跟着下一个信号。

b)初始条件是什么?

c)用这两个信号在 10 微秒内可以发送多少条不同的信息?

36. 一个汽车司机只用 5 美分和 10 美分硬币付过桥费，每次向收费机投一个硬币。

a)求与这个汽车司机付费 n 美分的不同方式数有关的递推关系(考虑使用硬币的次序)。

b)这个司机付费 45 美分有多少种可能的方式?

37. a)如果 R_n 是一个平面被 n 条直线划分出的区域个数，其中没有两条直线是平行的也没有 3 条直线交于一点，找出由 R_n 满足的递推关系。

b)使用迭代求出 R_n。

*38. a)找出由 R_n 满足的递推关系，其中 R_n 是一个球面被 n 个大圆(球面与通过球心的平面的交线)划分的区域个数，如果没有 3 个大圆交于一点。

b)使用迭代求出 R_n。

*39. a)找出由 S_n 满足的递推关系，其中 S_n 是三维空间被 n 个平面分成的区域数，如果每 3 个平面交于一点，但没有 4 个平面交于一点。

b)使用迭代求出 S_n。

40. 求出与具有偶数个0的n位二进位串个数有关的递推关系。

41. 包含偶数个0的7位二进位串有多少个？

42. a)找到与用1×2的多米诺牌完全覆盖$2\times n$的棋盘的方式数有关的递推关系。[提示：分别考虑对棋盘右上角的位置用一张多米诺牌水平放置和垂直放置的覆盖方式。]

    b)关于(a)中递推关系的初始条件是什么？

    c)用1×2的多米诺牌完全覆盖2×17的棋盘有多少种方式？

43. a)用地砖铺一条人行道，地砖是红色、绿色或灰色的。如果没有两块红砖相邻且同色的地砖是不加区别的，找出与用n块砖铺一条路的方式数有关的递推关系。

    b)对于(a)中的递推关系有什么初始条件？

    c)用7块砖铺一条在(a)中所描述的路有多少种方式？

44. 证明斐波那契数满足递推关系$f_n=5f_{n-4}+3f_{n-5}$，$n=5, 6, 7, \cdots$，其中递推关系具有初始条件$f_0=0$，$f_1=1$，$f_2=1$，$f_3=2$，$f_4=3$。用这个递推关系证明f_{5n}可被5整除，$n=1, 2, 3, \cdots$。

*45. 设$S(m, n)$表示从m元素集合到n元素集合的映上函数个数。证明$S(m, n)$满足递推关系

$$S(m,n)=n^m-\sum_{k=1}^{n-1}C(n,k)S(m,k)$$

其中$m\geqslant n$且$n>1$，初始条件是$S(m, 1)=1$。

46. a)写出为确定相乘次序而在乘积$x_0\cdot x_1\cdot x_2\cdot x_3\cdot x_4$中加括号的所有方式。

    b)使用在例8所建立的递推关系计算C_4，即为确定相乘的次序在5个数的乘积中加括号的方式数。验证你在(a)列出的方式数是正确的。

    c)使用在例8的解答中所提到的关于C_n的封闭公式，通过求出C_4检验你在(b)得到的结果。

47. a)使用在例8所建立的递推关系确定C_5，即为确定相乘的次序在6个数的乘积中加括号的方式数。

    b)使用在例8的解答中所提到的关于C_5的封闭公式检验你在(b)得到的结果。

*48. 在汉诺塔难题中，假设我们的目标是把所有的n个盘子从柱1移到柱3，但我们不能直接在柱1和柱3之间移动盘子。每次移动盘子必须通过柱2，并且，我们不能把较大的盘子放在较小的盘子上面。

    a)找出与求解这个具有附加限制条件的n个盘子的难题所需移动次数有关的递推关系。

    b)解这个递推关系来确定求解这个n个盘子难题所需移动次数的公式。

    c)有多少种不同的方法把n个盘子安排在3个柱子上使得没有一个较大的盘子放在较小的盘子上面？

    d)显示在这个变形难题的解中得到的对n个盘子的各种可能的安排。

练习49～53是格雷厄姆·克努斯(Graham Knuth)和帕塔什尼克(Patashnik)在[GrKnPa94]所描述的约瑟夫问题(Josephus Problem)的一种变形。这个问题来源于历史学家弗拉维乌斯·约瑟夫的一本账。41个犹太叛民在一世纪犹太-罗马战争期间被罗马人追赶逃入山洞，约瑟夫是这群人中的一个。这些叛民宁愿死也不愿被俘。他们决定围成一个圆圈并且围着这个圆圈重复数数，每数到3就杀掉这个位置的人而留下其他的人。但是约瑟夫和另一个叛民不愿意这样被杀掉。他们确定了应该站的位置是最后两个活下来的叛民的位置。我们考虑的问题开始时有n个人，记为1到n，站成一个圆圈。每一步，每第2个仍旧活着的人将被排除，直到只剩下一个人为止。我们把生还的人数记作$J(n)$。

49. 对每个正整数n的值，$1\leqslant n\leqslant16$，确定$J(n)$的值。

50. 使用你在练习49找到的值猜想一个关于$J(n)$的公式。[提示：写$n=2^m+k$，其中m是非负整数，k是小于2^m的非负整数。]

51. 对于$n\geqslant1$，证明$J(n)$满足递推关系$J(2n)=2J(n)-1$和$J(2n+1)=2J(n)+1$，且$J(1)=1$。

52. 用练习51的递推关系根据数学归纳法证明你在练习50所猜想的公式。

53. 根据你关于$J(n)$的公式确定$J(100)$，$J(1000)$和$J(10\,000)$。

练习54～61涉及雷夫难题，即具有4个柱和n个盘子的汉诺塔的变形问题。在给出这些练习之前，我们描述一个弗雷姆-斯图尔特(Frame-Stewart)算法，它把盘子从柱1移到柱4并且没有较大的盘子放在较

小的盘子上面。给定盘子数 n 作为输入，这个算法依赖于一个整数 k 的选择，$1\leqslant k\leqslant n$。当只有一个盘子时，把它从柱 1 移到柱 4，然后算法停止。对于 $n>1$，算法递归地使用下面的 3 步。首先使用所有的 4 根柱递归地把最小的 $n-k$ 个盘子从柱 1 移到柱 2。下一步使用汉诺塔问题的三根柱算法，不使用放 $n-k$ 个最小盘子的柱，把 k 个最大的盘子递归地从柱 1 移到柱 4。最后，使用所有 4 根柱递归地将 $n-k$ 个最小的盘子移到柱 4。弗雷姆和斯图尔特证明，使用他们的算法，为了达到最少的移动次数，应该选择 k 使得 n 是不超过第 k 个三角形数 $t_k=k(k+1)/2$ 的最小的正整数，即 $t_{k-1}<n\leqslant t_k$。有一个未被证实的猜想，称为*弗雷姆猜想*，就是不管盘子怎样移动，该算法对于求解这个难题所需要的移动次数最少。

54. 证明具有 3 个盘子的雷夫难题最少可以使用 5 次移动求解。
55. 证明具有 4 个盘子的雷夫难题最少可以使用 9 次移动求解。
56. 描述弗雷姆-斯图尔特算法所做的移动，并选择 k 使得在下面每种情况下所需要的移动次数最少。
    a)5 个盘子　　b)6 个盘子　　c)7 个盘子　　d)8 个盘子

*57. 证明如果 $R(n)$ 是由弗雷姆-斯图尔特算法求解具有 n 个盘子的雷夫难题所使用的移动次数，这里选择 k 是满足 $n\leqslant k(k+1)/2$ 的最小的整数，那么 $R(n)$ 满足递推关系 $R(n)=2R(n-k)+2^k-1$ 和 $R(0)=0$，$R(1)=1$。

*58. 证明如果 k 如练习 57 所选，那么 $R(n)-R(n-1)=2^{k-1}$。

*59. 证明如果 k 如练习 57 所选，那么 $R(n)=\sum_{i=1}^{k} i2^{i-1}-(t_k-n)2^{k-1}$。

*60. 用练习 59 给出对所有的整数 $n(1\leqslant n\leqslant 25)$，求解雷夫难题所需移动次数的上界。

*61. 证明 $R(n)$ 是 $O(\sqrt{n}2^{\sqrt{2n}})$。

设 $\{a_n\}$ 是实数序列，这个序列的*后向差分*递归地定义如下：
第一差分 ∇a_n 是

$$\nabla a_n=a_n-a_{n-1}$$

从 $\nabla^k a_n$ 得到*第 $k+1$ 差分* $\nabla^{k+1}a_n$，即
$\nabla^{k+1}a_n=\nabla^k a_n-\nabla^k a_{n-1}$

62. 求关于序列 $\{a_n\}$ 的 ∇a_n，其中
    a) $a_n=4$　　b) $a_n=2n$　　c) $a_n=n^2$　　d) $a_n=2^n$
63. 对于在练习 62 中的序列求 $\nabla^2 a_n$。
64. 证明 $a_{n-1}=a_n-\nabla a_n$。
65. 证明 $a_{n-2}=a_n-2\nabla a_n+\nabla^2 a_n$。

*66. 证明 a_{n-k} 可以用 a_n，∇a_n，$\nabla^2 a_n$，…，$\nabla^k a_n$ 的项表示。

67. 用 a_n，∇a_n，$\nabla^2 a_n$ 的项表示递推关系 $a_n=a_{n-1}+a_{n-2}$。
68. 证明关于序列 $\{a_n\}$ 的任何递推关系都可以用 a_n，∇a_n，$\nabla^2 a_n$，…的项表示。涉及这个序列和它的差分的等式叫做*差分方程*。

6.2 求解递推关系

6.2.1 引言

各种各样的递推关系出现在模型里。某些递推关系可以用迭代或者其他的特别技术求解。但是，有一类重要的递推关系可以用一种系统的方法明确地求解。在这种递推关系中，序列的项由它的前项的线性组合来表示。

定义 1 一个常系数的 k 阶线性齐次递推关系是形如

$$a_n=c_1a_{n-1}+c_2a_{n-2}+\cdots+c_ka_{n-k}$$

的递推关系，其中 c_1，c_2，…，c_k 是实数，$c_k\neq 0$。

这个定义中的递推关系是*线性的*，因为它的右边是序列前项的倍数之和。这个递推关系是

齐次的，因为所出现的各项都是 a_j 的倍数。序列各项的系数都是常数而不是依赖于 n 的函数。阶为 k 是因为 a_n 由序列前面的 k 项来表示。

根据数学归纳法第二原理，满足这个定义的递推关系的序列由这个递推关系和 k 个初始条件

$$a_0 = C_0, a_1 = C_1, \cdots, a_{k-1} = C_{k-1}$$

唯一地确定。

例 1 递推关系 $P_n=(1.11)P_{n-1}$ 是 1 阶的线性齐次递推关系。递推关系 $f_n=f_{n-1}+f_{n-2}$ 是 2 阶的线性齐次递推关系。递推关系 $a_n=a_{n-5}$ 是 5 阶的线性齐次递推关系。 ■

下面是一些常系数的但不是线性齐次递推关系的例子。

例 2 递推关系 $a_n=a_{n-1}+a_{n-2}^2$ 不是线性的。递推关系 $H_n=2H_{n-1}+1$ 不是齐次的。递推关系 $B_n=nB_{n-1}$ 不是常系数的。 ■

研究线性齐次递推关系有两个理由。第一，在建立问题的模型时经常出现这种递推关系。第二，它们可以用系统的方法求解。

6.2.2 求解常系数线性齐次递推关系

求解常系数线性齐次递推关系的基本方法是寻找形如 $a_n=r^n$ 的解，其中 r 是常数。注意 $a_n=r^n$ 是递推关系 $a_n=c_1a_{n-1}+c_2a_{n-2}+\cdots+c_ka_{n-k}$ 的解，当且仅当

$$r^n = c_1r^{n-1} + c_2r^{n-2} + \cdots + c_kr^{n-k}$$

当等式的两边除以 r^{n-k} 并且从左边减去右边时，我们得到等价的方程

$$r^k - c_1r^{k-1} - c_2r^{k-2} - \cdots - c_{k-1}r - c_k = 0$$

因此，序列 $\{a_n\}$ 以 $a_n=r^n$ 作为解，当且仅当 r 是这后一个方程的解。这个方程叫做该递推关系的*特征方程*。方程的解叫做该递推关系的*特征根*。正如我们将要看到的，可以用这些特征根给出这种递推关系的所有解的显式公式。

我们首先看一个 2 阶常系数线性齐次递推关系的处理结果。然后，叙述相应的阶可能大于 2 的一般性结果。由于得到一般性结果所需要的证明比较复杂，本书不再赘述。

我们现在回到 2 阶线性齐次递推关系。首先，考虑存在两个不等的特征根的情况。

定理 1 设 c_1 和 c_2 是实数。假设 $r^2-c_1r-c_2=0$ 有两个不等的根 r_1 和 r_2，那么序列 $\{a_n\}$ 是递推关系 $a_n=c_1a_{n-1}+c_2a_{n-2}$ 的解，当且仅当 $a_n=\alpha_1r_1^n+\alpha_2r_2^n$，$n=0，1，2，\cdots$，其中 α_1 和 α_2 是常数。

证 证明这个定理必须做两件事。首先，必须证明如果 r_1 和 r_2 是特征方程的根，并且 α_1 和 α_2 是常数，那么序列 $\{a_n\}$ ($a_n=\alpha_1r_1^n+\alpha_2r_2^n$) 是递推关系的解。第二，必须证明如果序列 $\{a_n\}$ 是解，那么对某个常数 α_1 和 α_2 有 $a_n=\alpha_1r_1^n+\alpha_2r_2^n$。

现在我们将证明如果 $a_n=\alpha_1r_1^n+\alpha_2r_2^n$，那么序列 $\{a_n\}$ 是递推关系的解。因为 r_1 和 r_2 是 $r^2-c_1r-c_2=0$ 的根，从而 $r_1^2=c_1r_1+c_2$，$r_2^2=c_1r_2+c_2$。

从这些等式可以看出

$$\begin{aligned}
c_1a_{n-1}+c_2a_{n-2} &= c_1(\alpha_1r_1^{n-1}+\alpha_2r_2^{n-1})+c_2(\alpha_1r_1^{n-2}+\alpha_2r_2^{n-2}) \\
&= \alpha_1r_1^{n-2}(c_1r_1+c_2)+\alpha_2r_2^{n-2}(c_1r_2+c_2) \\
&= \alpha_1r_1^{n-2}r_1^2+\alpha_2r_2^{n-2}r_2^2 \\
&= \alpha_1r_1^n+\alpha_2r_2^n \\
&= a_n
\end{aligned}$$

这证明了序列 $\{a_n\}$ 以 $a_n=\alpha_1r_1^n+\alpha_2r_2^n$ 作为递推关系的解。

为证明递推关系 $a_n=c_1a_{n-1}+c_2a_{n-2}$ 的每一个解 $\{a_n\}$ 都有形式 $a_n=\alpha_1r_1^n+\alpha_2r_2^n$，$n=0$，1，2，…，$\alpha_1$ 和 α_2 为某个常数。假设 $\{a_n\}$ 是递推关系的解，初始条件是 $a_0=C_0$，$a_1=C_1$。下面证明存在常数 α_1 和 α_2 使得具有 $a_n=\alpha_1r_1^n+\alpha_2r_2^n$ 的序列 $\{a_n\}$ 满足同样的初始条件。这要求

$$a_0=C_0=\alpha_1+\alpha_2$$
$$a_1=C_1=\alpha_1r_1+\alpha_2r_2$$

我们可以求解这两个关于 α_1 和 α_2 的方程。从第一个方程得到 $\alpha_2=C_0-\alpha_1$。把它代入第二个方程得

$$C_1=\alpha_1r_1+(C_0-\alpha_1)r_2$$

因此，

$$C_1=\alpha_1(r_1-r_2)+C_0r_2$$

这说明了

$$\alpha_1=\frac{(C_1-C_0r_2)}{r_1-r_2}$$

和

$$\alpha_2=C_0-\alpha_1=C_0-\frac{(C_1-C_0r_2)}{r_1-r_2}=\frac{C_0r_1-C_1}{r_1-r_2}$$

这里关于 α_1 和 α_2 的表达式依赖于 $r_1\neq r_2$ 的事实。(当 $r_1=r_2$ 时，这个定理不成立。)因此，由于这两个 α_1 和 α_2 的值，具有 $\alpha_1r_1^n+\alpha_2r_2^n$ 的序列 $\{a_n\}$ 满足这两个初始条件。因为这个递推关系和这些初始条件唯一地确定了这个序列，从而得出 $a_n=\alpha_1r_1^n+\alpha_2r_2^n$。 □

常系数线性齐次递推关系的特征根可能是复数。定理 1(和本节后面的定理)在这种情况下仍旧适用。具有复数特征根的递推关系在本书中不进行讨论。熟悉复数的读者可以做节末的练习 38 和 39。

例 3 和例 4 说明怎样用定理 1 求解递推关系。

例 3 什么是如下递推关系的解?

$$a_n=a_{n-1}+2a_{n-2}$$

其中 $a_0=2$ 和 $a_1=7$。

解 可用定理 1 求解这个问题。递推关系的特征方程是 $r^2-r-2=0$。它的根是 $r=2$ 和 $r=-1$。因此，序列 $\{a_n\}$ 是递推关系的解当且仅当

$$a_n=\alpha_12^n+\alpha_2(-1)^n$$

α_1 和 α_2 是常数。由初始条件，得

$$a_0=2=\alpha_1+\alpha_2$$
$$a_1=7=\alpha_1\cdot 2+\alpha_2\cdot(-1)$$

求解这两个等式得 $\alpha_1=3$ 和 $\alpha_2=-1$。因此，关于这个递推关系和初始条件的解是序列 $\{a_n\}$，其中

$$a_n=3\cdot 2^n-(-1)^n$$ ■

例 4 找一个关于斐波那契数的显式公式。

解 斐波那契数的序列满足递推关系 $f_n=f_{n-1}+f_{n-2}$ 和初始条件 $f_0=0$ 及 $f_1=1$。特征方程 $r^2-r-1=0$ 的根是 $r_1=(1+\sqrt{5})/2$ 和 $r_2=(1-\sqrt{5})/2$。因此，从定理 1 得到斐波那契数由

$$f_n=\alpha_1\left(\frac{1+\sqrt{5}}{2}\right)^n+\alpha_2\left(\frac{1-\sqrt{5}}{2}\right)^n$$

给出，其中 α_1 和 α_2 为常数。可用初始条件 $f_0=0$ 和 $f_1=1$ 确定这些常数。我们有

$$f_0 = \alpha_1 + \alpha_2 = 0$$

$$f_1 = \alpha_1\left(\frac{1+\sqrt{5}}{2}\right) + \alpha_2\left(\frac{1-\sqrt{5}}{2}\right) = 1$$

对这些关于 α_1 和 α_2 的联立方程的解是

$$\alpha_1 = 1/\sqrt{5}, \quad \alpha_2 = -1/\sqrt{5}$$

于是，斐波那契数由下面的式子给出：

$$f_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$$ ■

当存在二重特征根时定理 1 不再适用。如果发生这种情况，当 r_0 是特征方程的一个二重根时那么 $a_n = nr_0^n$ 是递推关系的另一个解。定理 2 说明了怎样处理这种情况。

定理 2 设 c_1 和 c_2 是实数，$c_2 \neq 0$。假设 $r^2 - c_1 r - c_2 = 0$ 只有一个根 r_0。序列$\{a_n\}$是递推关系 $a_n = c_1 a_{n-1} + c_2 a_{n-2}$ 的解，当且仅当 $a_n = \alpha_1 r_0^n + \alpha_2 n r_0^n$，$n = 0, 1, 2, \cdots$，其中 α_1 和 α_2 是常数。

定理 2 的证明留作节末的练习。例 5 说明了这个定理的应用。

例 5 具有初始条件 $a_0 = 1$ 和 $a_1 = 6$ 的递推关系

$$a_n = 6a_{n-1} - 9a_{n-2}$$

的解是什么？

解 $r^2 - 6r + 9 = 0$ 唯一的根是 $r = 3$。因此，这个递推关系的解是：

$$a_n = \alpha_1 3^n + \alpha_2 n 3^n$$

其中 α_1 和 α_2 是常数。使用初始条件得

$$a_0 = 1 = \alpha_1$$

$$a_1 = 6 = \alpha_1 \cdot 3 + \alpha_2 \cdot 3$$

求解这两个方程得 $\alpha_1 = 1$ 和 $\alpha_2 = 1$。从而，这个具有给定初始条件的递推关系的解是

$$a_n = 3^n + n3^n$$ ■

我们现在叙述这个关于常系数线性齐次递推关系的解的一般性结果，这里的阶可以大于 2 且假定特征方程有不等的根。这个结果的证明留给读者作为练习。

定理 3 设 $c_1, c_2, \cdots, c_k$ 是实数。假设特征方程

$$r^k - c_1 r^{k-1} - \cdots - c_k = 0$$

有 k 个不等的根 $r_1, r_2, \cdots, r_k$。那么序列$\{a_n\}$是递推关系

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k}$$

的解，当且仅当

$$a_n = \alpha_1 r_1^n + \alpha_2 r_2^n + \cdots + \alpha_k r_k^n$$

$n = 0, 1, 2, \cdots$，其中 $\alpha_1, \alpha_2, \cdots, \alpha_k$ 是常数。

我们用例 6 说明定理的使用。

例 6 求出具有初始条件 $a_0 = 2$，$a_1 = 5$ 和 $a_2 = 15$ 的递推关系

$$a_n = 6a_{n-1} - 11a_{n-2} + 6a_{n-3}$$

的解。

解 这个递推关系的特征多项式是

$$r^3 - 6r^2 + 11r - 6$$

因为 $r^3 - 6r^2 + 11r - 6 = (r-1)(r-2)(r-3)$，所以特征根是 $r=1$，$r=2$ 和 $r=3$。因此，递推关

系的解的形式是

$$a_n = \alpha_1 \cdot 1^n + \alpha_2 \cdot 2^n + \alpha_3 \cdot 3^n$$

为找到常数 α_1，α_2 以及 α_3，使用初始条件得

$$a_0 = 2 = \alpha_1 + \alpha_2 + \alpha_3$$
$$a_1 = 5 = \alpha_1 + \alpha_2 \cdot 2 + \alpha_3 \cdot 3$$
$$a_2 = 15 = \alpha_1 + \alpha_2 \cdot 4 + \alpha_3 \cdot 9$$

当求解这三个关于 α_1，α_2 和 α_3 的联立方程时，得到 $\alpha_1=1$，$\alpha_2=-1$ 且 $\alpha_3=2$。于是，这个递推关系和给定初始条件的唯一解是满足

$$a_n = 1 - 2^n + 2 \cdot 3^n$$

的序列 $\{a_n\}$。 ■

我们现在叙述关于常系数线性齐次递推关系的最一般化的结果，这里允许特征方程有重根。要点是对于特征方程的每个根 r，通解是形如 $P(n)r^n$ 的项之和，其中 $P(n)$ 是 $m-1$ 次多项式，而 m 是这个根的重数。我们把证明作为一个挑战性的练习留给读者。

定理 4 设 c_1，c_2，…，c_k 是实数，假设特征方程

$$r^k - c_1 r^{k-1} - \cdots - c_k = 0$$

有 t 个不等的根 r_1，r_2，…，r_t，其重数分别为 m_1，m_2，…，m_t，满足 $m_i \geqslant 1$，$i=1$，2，…，t，且 $m_1+m_2+\cdots+m_t=k$。那么序列 $\{a_n\}$ 是递推关系

$$a_n = c_1 a_{n-1} + c_2 a_{n-2} + \cdots + c_k a_{n-k}$$

的解，当且仅当

$$\begin{aligned} a_n = &(\alpha_{1,0} + \alpha_{1,1} n + \cdots + \alpha_{1,m_1-1} n^{m_1-1}) r_1^n \\ &+ (\alpha_{2,0} + \alpha_{2,1} n + \cdots + \alpha_{2,m_2-1} n^{m_2-1}) r_2^n \\ &+ \cdots + (\alpha_{t,0} + \alpha_{t,1} n + \cdots + \alpha_{t,m_t-1} n^{m_t-1}) r_t^n \end{aligned}$$

$n=0$，1，2，…，其中 $\alpha_{i,j}$ 是常数，$1 \leqslant i \leqslant t$ 且 $0 \leqslant j \leqslant m_i - 1$。

例 7 说明在特征方程有重根时怎样用定理 4 求一个线性齐次递推关系的通解形式。

例 7 假设线性齐次递推关系的特征方程的根是 2，2，2，5，5 和 9(即有 3 个根，根 2 的重数为 3，根 5 的重数为 2，根 9 的重数为 1)。那么通解形式是什么？

解 由定理 4，解的一般形式是

$$(\alpha_{1,0} + \alpha_{1,1} n + \alpha_{1,2} n^2) 2^n + (\alpha_{2,0} + \alpha_{2,1} n) 5^n + \alpha_{3,0} 9^n$$

■

我们现在说明在特征方程有 3 重根时如何用定理 4 求解常系数线性齐次递推关系。

例 8 找出具有初始条件 $a_0=1$，$a_1=-2$ 和 $a_2=-1$ 的递推关系

$$a_n = -3a_{n-1} - 3a_{n-2} - a_{n-3}$$

的解。

解 这个递推关系的特征方程是

$$r^3 + 3r^2 + 3r + 1 = 0$$

因为 $r^3+3r^2+3r+1=(r+1)^3$，特征方程只有一个 3 重根 $r=-1$。由定理 4，这个递推关系的解是下述形式

$$a_n = \alpha_{1,0}(-1)^n + \alpha_{1,1} n(-1)^n + \alpha_{1,2} n^2 (-1)^n$$

为求出常数 $\alpha_{1,0}$，$\alpha_{1,1}$ 和 $\alpha_{1,2}$，使用初始条件，得到

$$a_0 = 1 = \alpha_{1,0}$$
$$a_1 = -2 = -\alpha_{1,0} - \alpha_{1,1} - \alpha_{1,2}$$

$$a_2 = -1 = \alpha_{1,0} + 2\alpha_{1,1} + 4\alpha_{1,2}$$

这3个方程的联立解是 $\alpha_{1,0}=1$，$\alpha_{1,1}=3$，$\alpha_{1,2}=-2$。于是，这个递推关系和给定初始条件的唯一解是序列 $\{a_n\}$，其中

$$a_n = (1+3n-2n^2)(-1)^n$$

■

6.2.3 常系数线性非齐次的递推关系

我们已经知道如何求解常系数线性齐次的递推关系。是否有一种相对简单的技术来求解如像 $a_n=3a_{n-1}+2n$ 的常系数线性但是非齐次的递推关系呢？我们将看到，仅仅对某些特定类型的递推关系存在肯定的回答。

递推关系 $a_n=3a_{n-1}+2n$ 是一个**常系数线性非齐次递推关系**，即形如

$$a_n = c_1a_{n-1} + c_2a_{n-2} + \cdots + c_ka_{n-k} + F(n)$$

的递推关系的例子，其中 c_1，c_2，…，c_k 是实数，$F(n)$是只依赖于 n 且不恒为0的函数。递推关系

$$a_n = c_1a_{n-1} + c_2a_{n-2} + \cdots + c_ka_{n-k}$$

叫做**相伴的齐次递推关系**。它在非齐次递推关系的解中起了重要的作用。

例9 递推关系 $a_n=a_{n-1}+2^n$，$a_n=a_{n-1}+a_{n-2}+n^2+n+1$，$a_n=3a_{n-1}+n3^n$ 和 $a_n=a_{n-1}+a_{n-2}+a_{n-3}+n!$ 是常系数线性非齐次递推关系。相伴的线性齐次递推关系分别是 $a_n=a_{n-1}$，$a_n=a_{n-1}+a_{n-2}$，$a_n=3a_{n-1}$ 和 $a_n=a_{n-1}+a_{n-2}+a_{n-3}$。 ■

关于常系数线性非齐次递推关系的一个关键事实是，每个解都是一个特解与相伴的线性齐次递推关系的一个解之和，正如定理5所述。

定理5 如果 $\{a_n^{(p)}\}$ 是常系数非齐次线性递推关系

$$a_n = c_1a_{n-1} + c_2a_{n-2} + \cdots + c_ka_{n-k} + F(n)$$

的一个特解，那么每个解都是 $\{a_n^{(p)}+a_n^{(h)}\}$ 的形式，其中 $\{a_n^{(h)}\}$ 是相伴的齐次递推关系

$$a_n = c_1a_{n-1} + c_2a_{n-2} + \cdots + c_ka_{n-k}$$

的一个解。

证 由于 $\{a_n^{(p)}\}$ 是非齐次递推关系的特解，我们知道

$$a_n^{(p)} = c_1a_{n-1}^{(p)} + c_2a_{n-2}^{(p)} + \cdots + c_ka_{n-k}^{(p)} + F(n)$$

现在假设 $\{b_n\}$ 是常系数非齐次递推关系的第二个解，使得

$$b_n = c_1b_{n-1} + c_2b_{n-2} + \cdots + c_kb_{n-k} + F(n)$$

从第二个等式减去第一个等式得

$$b_n - a_n^{(p)} = c_1(b_{n-1} - a_{n-1}^{(p)}) + c_2(b_{n-2} - a_{n-2}^{(p)}) + \cdots + c_k(b_{n-k} - a_{n-k}^{(p)})$$

从而得到 $\{b_n-a_n^{(p)}\}$ 是相伴的线性齐次递推关系的一个解，比如说是 $\{a_n^{(h)}\}$。因此，对所有的 n 有 $b_n=a_n^{(p)}+a_n^{(h)}$。 □

由定理5，我们看到求解常系数非齐次递推关系的关键是找一个特解。然后每个解都是这个特解与相伴的齐次递推关系的一个解之和。尽管不存在对每个函数 $F(n)$ 都有效的一般性方法来求这种解，但有某些技术对特定的函数类型 $F(n)$，例如多项式函数与常数的幂函数有效。例10和例11就说明了这一点。

例10 求递推关系 $a_n=3a_{n-1}+2n$ 的所有的解。具有 $a_1=3$ 的解是什么？

解 为求解这个常系数线性非齐次递推关系，我们需要求解与它相伴的线性齐次方程并且找到一个关于给定非齐次方程的特解。相伴的线性齐次方程是 $a_n=3a_{n-1}$。它的解是 $a_n^{(h)}=\alpha 3^n$，

其中 α 是常数。

我们现在找一个特解。因为 $F(n)=2n$ 是 n 的 1 次多项式，解的一个合理的尝试就是 n 的线性函数，比如说 $p_n=cn+d$，其中 c 和 d 是常数。为确定是否存在这种形式的解，假设 $p_n=cn+d$ 是一个这样的解。那么方程 $a_n=3a_{n-1}+2n$ 就变成 $cn+d=3(c(n-1)+d)+2n$。化简和归并同类项得 $(2+2c)n+(2d-3c)=0$。从而，$cn+d$ 是一个解，当且仅当 $2+2c=0$ 和 $2d-3c=0$。这说明 $cn+d$ 是一个解当且仅当 $c=-1$ 和 $d=-3/2$。因而，$a_n^{(p)}=-n-3/2$ 是一个特解。

根据定理 5，所有的解都是下述形式

$$a_n = a_n^{(p)} + a_n^{(h)} = -n-3/2+\alpha\cdot 3^n$$

其中 α 是常数。

为找出具有 $a_1=3$ 的解，在得到的通解公式中令 $n=1$。我们有 $3=-1-3/2+3\alpha$，这就推出 $\alpha=11/6$。我们要找的解是 $a_n=-n-3/2+(11/6)3^n$。■

例 11 求出下述递推关系

$$a_n = 5a_{n-1}-6a_{n-2}+7^n$$

的所有的解。

解 这是一个线性非齐次递推关系。它的相伴的齐次递推关系

$$a_n = 5a_{n-1}-6a_{n-2}$$

的解是 $a_n^{(h)}=\alpha_1\cdot 3^n+\alpha_2\cdot 2^n$，其中 α_1 和 α_2 是常数。因为 $F(n)=7^n$，一个合理的解是 $a_n^{(p)}=C\cdot 7^n$，其中 C 是常数。把这些项代入递推关系得 $C\cdot 7^n=5C\cdot 7^{n-1}-6C\cdot 7^{n-2}+7^n$。提出公因式 7^{n-2}，这个等式变成 $49C=35C-6C+49$，从而推出 $20C=49$ 或 $C=49/20$。于是，$a_n^{(p)}=(49/20)7^n$ 是特解。由定理 5，所有的解都有下述形式

$$a_n = \alpha_1\cdot 3^n+\alpha_2\cdot 2^n+(49/20)7^n$$

■

在例 10 和例 11 中，我们凭经验猜想了一个特定形式的解。在两种情况下，我们都能找到特解。这并不是偶然的。每当 $F(n)$ 是 n 的多项式和一个常数的 n 次幂之积时，我们就能知道一个特解恰好是什么形式，正如定理 6 所述。定理 6 的证明作为一个挑战性的练习留给读者。

定理 6 假设 $\{a_n\}$ 满足线性非齐次递推关系

$$a_n = c_1a_{n-1}+c_2a_{n-2}+\cdots+c_ka_{n-k}+F(n)$$

其中 c_1，c_2，…，c_k 是实数，且

$$F(n) = (b_tn^t+b_{t-1}n^{t-1}+\cdots+b_1n+b_0)s^n$$

其中 b_0，b_1，…，b_t 和 s 是实数。当 s 不是相伴的线性齐次递推关系的特征方程的根时，存在一个下述形式的特解

$$(p_tn^t+p_{t-1}n^{t-1}+\cdots+p_1n+p_0)s^n$$

当 s 是特征方程的根且它的重数是 m 时，存在一个下述形式的特解

$$n^m(p_tn^t+p_{t-1}n^{t-1}+\cdots+p_1n+p_0)s^n$$

注意当 s 是相伴的线性齐次递推关系的特征方程的 m 重根时，因式 n^m 保证给出的特解不是相伴的线性齐次递推关系的一个解。我们下面给出一个例子说明定理 6 所提供的特解形式。

例 12 当 $F(n)=3^n$，$F(n)=n3^n$，$F(n)=n^22^n$ 和 $F(n)=(n^2+1)3^n$ 时，线性非齐次递推关系 $a_n=6a_{n-1}-9a_{n-2}+F(n)$ 的特解有什么形式？

解 相伴的线性齐次递推关系是 $a_n=6a_{n-1}-9a_{n-2}$。它的特征方程 $r^2-6r+9=(r-3)^2=0$ 有一个 2 重的单根 3。$F(n)$ 的形式为 $P(n)s^n$，其中 $P(n)$ 是一个多项式，s 是一个常数。为应用定理 6，我们需要知道 s 是否是这个特征方程的根。

由于 $s=3$ 是重数 $m=2$ 的根而 $s=2$ 不是根，定理 6 告诉我们如果 $F(n)=3^n$，特解的形式是

$p_0n^2 3^n$；如果 $F(n)=n3^n$，特解的形式是 $n^2(p_1n+p_0)3^n$；如果 $F(n)=n^2 2^n$，特解的形式是 $(p_2n^2+p_1n+p_0)2^n$；如果 $F(n)=(n^2+1)3^n$，特解的形式是 $n^2(p_2n^2+p_1n+p_0)3^n$。 ■

在求解定理 6 所谈的那种类型的递推关系时，若 $s=1$ 一定要小心处理。特别是把定理用于 $F(n)=b_tn^t+b_{t-1}n^{t-1}+\cdots+b_1n+b_0$，参数 s 取值 $s=1$ 时（尽管项 1^n 没有明确地出现）的情况。根据这个定理，解的形式就依赖于是否 1 是相伴的线性齐次递推关系的特征方程的根。这将在例 13 中说明，它说明了怎样用定理 6 找出前 n 个正整数之和的公式。

例 13 设 a_n 是前 n 个正整数的和，即

$$a_n = \sum_{k=1}^{n} k$$

注意，a_n 满足线性非齐次递推关系

$$a_n = a_{n-1} + n$$

（为从前 $n-1$ 个正整数的和 a_{n-1} 得到前 n 个正整数的和 a_n，只需加上 n 即可）。注意初始条件是 $a_1=1$。

对于 a_n，相伴的线性齐次递推关系是

$$a_n = a_{n-1}$$

这个齐次递推关系的解是 $a_n^{(h)}=c(1)^n=c$，其中 c 是一个常数。为找到 $a_n=a_{n-1}+n$ 的所有的解，我们仅需要找一个特解。由定理 6，由于 $F(n)=n=n(1)^n$ 且 $s=1$ 是相伴的线性齐次递推关系的特征方程的 1 阶根，存在一个形如 $n(p_1n+p_0)=p_1n^2+p_0n$ 的特解。

把它代入递推关系得到 $p_1n^2+p_0n=p_1(n-1)^2+p_0(n-1)+n$。化简得 $n(2p_1-1)+(p_0-p_1)=0$，这意味着 $2p_1-1=0$ 和 $p_0-p_1=0$，即 $p_0=p_1=1/2$。因此

$$a_n^{(p)} = \frac{n^2}{2} + \frac{n}{2} = \frac{n(n+1)}{2}$$

是一个特解。所以，原递推关系 $a_n=a_{n-1}+n$ 的所有的解由 $a_n=a_n^{(h)}+a_n^{(p)}=c+n(n+1)/2$ 给出。由于 $a_1=1$，我们有 $1=a_1=c+1\cdot 2/2=c+1$，故 $c=0$。从而 $a_n=n(n+1)/2$。（这和 3.2 节表 3-2 给出的以及前面导出的公式一样。） ■

练习

1. 确定下面哪些是常系数线性齐次递推关系，如果是，求它们的阶。
   a) $a_n=3a_{n-1}+4a_{n-2}+5a_{n-3}$
   b) $a_n=2na_{n-1}+a_{n-2}$
   c) $a_n=a_{n-1}+a_{n-4}$
   d) $a_n=a_{n-1}+2$
   e) $a_n=a_{n-1}^2+a_{n-2}$
   f) $a_n=a_{n-2}$
   g) $a_n=a_{n-1}+n$
2. 确定下面哪些是常系数线性齐次递推关系，如果是，求它们的阶。
   a) $a_n=3a_{n-2}$
   b) $a_n=3$
   c) $a_n=a_{n-1}^2$
   d) $a_n=a_{n-1}+2a_{n-3}$
   e) $a_n=a_{n-1}/n$
   f) $a_n=a_{n-1}+a_{n-2}+n+3$
   g) $a_n=4a_{n-2}+5a_{n-4}+9a_{n-7}$
3. 求解下述具有给定初始条件的递推关系
   a) $a_n=2a_{n-1}$，$n\geqslant 1$，$a_0=3$
   b) $a_n=a_{n-1}$，$n\geqslant 1$，$a_0=2$
   c) $a_n=5a_{n-1}-6a_{n-2}$，$n\geqslant 2$，$a_0=1$，$a_1=0$
   d) $a_n=4a_{n-1}-4a_{n-2}$，$n\geqslant 2$，$a_0=6$，$a_1=8$
   e) $a_n=-4a_{n-1}-4a_{n-2}$，$n\geqslant 2$，$a_0=0$，$a_1=1$
   f) $a_n=4a_{n-2}$，$n\geqslant 2$，$a_0=0$，$a_1=4$
   g) $a_n=a_{n-2}/4$，$n\geqslant 2$，$a_0=1$，$a_1=0$
4. 求解下述具有给定初始条件的递推关系
   a) $a_n=a_{n-1}+6a_{n-2}$，$n\geqslant 2$，$a_0=3$，$a_1=6$
   b) $a_n=7a_{n-1}-10a_{n-2}$，$n\geqslant 2$，$a_0=2$，$a_1=1$

c) $a_n=6a_{n-1}-8a_{n-2}$，$n\geqslant 2$，$a_0=4$，$a_1=10$

d) $a_n=2a_{n-1}-a_{n-2}$，$n\geqslant 2$，$a_0=4$，$a_1=1$

e) $a_n=a_{n-2}$，$n\geqslant 2$，$a_0=5$，$a_1=-1$

f) $a_n=-6a_{n-1}-9a_{n-2}$，$n\geqslant 2$，$a_0=3$，$a_1=-3$

g) $a_{n+2}=-4a_{n+1}+5a_n$，$n\geqslant 0$，$a_0=2$，$a_1=8$

5. 使用6.1节练习35描述的两个信号在 n 微秒内可以传送多少不同的信息？
6. 如果传送1个信号要1微秒，传送另外2个信号中的每一个都需要2微秒，且在信息中一个信号紧接着下一个信号，使用这3个不同的信号在 n 微秒内可以传送多少个不同的信息？
7. 使用1×2和2×2的块铺满一块2×n的长方形板，有多少种方式？
8. 一个关于每年捕龙虾数的模型基于如下的假设：1年捕捞的龙虾数是前2年捕捞龙虾数的平均值。

   a) 找出一个关于 $\{L_n\}$ 的递推关系，其中 L_n 是在这个模型的假设下第 n 年捕捞的龙虾数。

   b) 如果在第1年捕捞了100 000只龙虾且第2年捕捞了300 000只龙虾，求 L_n。
9. 年初把一笔100 000美元的钱存入一个投资基金。在每年的最后一天得到两份红利。第一份红利是当年账上钱数的20%。第二份红利是前一年账上钱数的45%。

   a) 如果不允许取钱，找出一个关于 $\{P_n\}$ 的递推关系，其中 P_n 是第 n 年末账上的钱数。

   b) 如果不允许取钱，n 年以后账上有多少钱？

*10. 证明定理2。

11. 卢卡斯数满足递推关系

$$L_n = L_{n-1} + L_{n-2}$$

和初始条件 $L_0=2$ 和 $L_1=1$。

   a) 证明 $L_n=f_{n-1}+f_{n+1}$，$n=2, 3, \cdots$，其中 f_n 是第 n 个斐波那契数。

   b) 求出卢卡斯数的显式公式。

12. 求解 $a_n=2a_{n-1}+a_{n-2}-2a_{n-3}$，$n=3, 4, 5, \cdots$，且 $a_0=3$，$a_1=6$，$a_2=0$。
13. 求解 $a_n=7a_{n-2}+6a_{n-3}$，$a_0=9$，$a_1=10$，$a_2=32$。
14. 求解 $a_n=5a_{n-2}-4a_{n-4}$，$a_0=3$，$a_1=2$，$a_2=6$，$a_3=8$。
15. 求解 $a_n=2a_{n-1}+5a_{n-2}-6a_{n-3}$，$a_0=7$，$a_1=-4$，$a_2=8$。

*16. 证明定理3。

17. 证明下述涉及斐波那契数和二项式系数的恒等式：

$$f_{n+1} = C(n,0) + C(n-1,1) + \cdots + C(n-k,k)$$

其中 n 是正整数且 $k=\lceil n/2\rceil$。[提示：设 $a_n=C(n, 0)+C(n-1, 1)+\cdots+C(n-k, k)$。证明序列 $\{a_n\}$ 和斐波那契序列满足的递推关系和初始条件一样。]

18. 求解递推关系 $a_n=6a_{n-1}-12a_{n-2}+8a_{n-3}$，$a_0=-5$，$a_1=4$，$a_2=88$。
19. 求解递推关系 $a_n=-3a_{n-1}-3a_{n-2}-a_{n-3}$，$a_0=5$，$a_1=-9$，$a_2=15$。
20. 找出递推关系 $a_n=8a_{n-2}-16a_{n-4}$ 的解的一般形式。
21. 如果线性齐次递推关系的特征方程的根是1，1，1，1，−2，−2，−2，3，3，−4，那么它的解的一般形式是什么？
22. 如果线性齐次递推关系的特征方程的根是−1，−1，−1，2，2，5，5，7，那么它的解的一般形式是什么？
23. 考虑非齐次线性递推关系 $a_n=3a_{n-1}+2^n$。

   a) 证明 $a_n=-2^{n+1}$ 是这个递推关系的一个解。

   b) 使用定理5找出这个递推关系的所有的解。

   c) 找出具有 $a_0=1$ 的解。
24. 考虑非齐次线性递推关系 $a_n=2a_{n-1}+2^n$。

   a) 证明 $a_n=n2^n$ 是这个递推关系的一个解。

   b) 使用定理5找出这个递推关系的所有的解。

   c) 找出具有 $a_0=2$ 的解。

25. a)确定常数 A 和 B 的值，使得 $a_n = An + B$ 是递推关系 $a_n = 2a_{n-1} + n + 5$ 的一个解。

 b)使用定理 5 找出这个递推关系的所有的解。

 c)找出这个递推关系具有 $a_0 = 4$ 的解。

26. 什么是线性非齐次递推关系 $a_n = 6a_{n-1} - 12a_{n-2} + 8a_{n-3} + F(n)$ 的特解的一般形式？如果

 a) $F(n) = n^2$ b) $F(n) = 2^n$ c) $F(n) = n2^n$ d) $F(n) = (-2)^n$

 e) $F(n) = n^2 2^n$ f) $F(n) = n^3(-2)^n$ g) $F(n) = 3$

27. 什么是线性非齐次递推关系 $a_n = 8a_{n-2} - 16a_{n-4} + F(n)$ 的特解的一般形式？如果

 a) $F(n) = n^3$ b) $F(n) = (-2)^n$ c) $F(n) = n2^n$ d) $F(n) = n^2 4^n$

 e) $F(n) = (n^2 - 2)(-2)^n$ f) $F(n) = n^4 2^n$ g) $F(n) = 2$

28. a)找出递推关系 $a_n = 2a_{n-1} + 2n^2$ 的所有的解。

 b)找出(a)中的递推关系具有初始条件 $a_1 = 4$ 的解。

29. a)找出递推关系 $a_n = 2a_{n-1} + 3^n$ 的所有的解。

 b)找出(a)中的递推关系具有初始条件 $a_1 = 5$ 的解。

30. a)找出递推关系 $a_n = -5a_{n-1} - 6a_{n-2} + 42 \cdot 4^n$ 的所有的解。

 b)找出这个递推关系具有初始条件 $a_1 = 56$ 和 $a_2 = 278$ 的解。

31. 找出递推关系 $a_n = 5a_{n-1} - 6a_{n-2} + 2^n + 3n$ 的所有的解。[提示：找形如 $qn2^n + p_1 n + p_2$ 的特解，其中 q，p_1，p_2 是常数。]

32. 找出递推关系 $a_n = 2a_{n-1} + 3 \cdot 2^n$ 的解。

33. 找出递推关系 $a_n = 4a_{n-1} - 4a_{n-2} + (n+1)2^n$ 的所有的解。

34. 找出递推关系 $a_n = 7a_{n-1} - 16a_{n-2} + 12a_{n-3} + n4^n$ 的具有 $a_0 = -2$，$a_1 = 0$，$a_2 = 5$ 的所有的解。

35. 找出递推关系 $a_n = 4a_{n-1} - 3a_{n-2} + 2^n + n + 3$ 的具有 $a_0 = 1$，$a_1 = 4$ 的解。

36. 设 a_n 是前 n 个完全平方的和，即 $a_n = \sum_{k=1}^{n} k^2$，证明序列 $\{a_n\}$ 满足线性非齐次递推关系 $a_n = a_{n-1} + n^2$ 和初始条件 $a_1 = 1$。使用定理 6 求解这个递推关系以确定关于 a_n 的公式。

37. 设 a_n 是前 n 个三角形数的和，即 $a_n = \sum_{k=1}^{n} t_k$，其中 $t_k = k(k+1)/2$。证明 $\{a_n\}$ 满足线性非齐次递推关系 $a_n = a_{n-1} + n(n+1)/2$ 和初始条件 $a_1 = 1$。使用定理 6 求解这个递推关系以确定关于 a_n 的公式。

38. a)求线性齐次递推关系 $a_n = 2a_{n-1} - 2a_{n-2}$ 的特征根。(注：这些根是复数。)

 b)求(a)的递推关系具有 $a_0 = 1$ 和 $a_1 = 2$ 的解。

*39. a)求线性齐次递推关系 $a_n = a_{n-4}$ 的特征根。(注：这些根包含复数。)

 b)求(a)的递推关系具有 $a_0 = 1$，$a_1 = 0$，$a_2 = -1$ 和 $a_3 = 1$ 的解。

*40. 求解联立递推关系

 $a_n = 3a_{n-1} + 2b_{n-1}$

 $b_n = a_{n-1} + 2b_{n-1}$

 初始条件 $a_0 = 1$ 和 $b_0 = 2$。

*41. a)用例 4 找到的关于第 n 个斐波那契数 f_n 的公式证明 f_n 是最接近

$$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n$$

 的整数。

 b)确定对哪些 n 有 f_n 大于

$$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n$$

 对哪些 n 有 f_n 小于

$$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n$$

42. 证明如果 $a_n = a_{n-1} + a_{n-2}$，$a_0 = s$ 和 $a_1 = t$，其中 s 和 t 是常数，那么对所有的正整数 n 有 $a_n = sf_{n-1} + tf_n$。
43. 用斐波那契数的项表示线性非齐次递推关系 $a_n = a_{n-1} + a_{n-2} + 1$ 的解，其中 $n \geqslant 2$，$a_0 = 0$，$a_1 = 1$。[提示：令 $b_n = a_n + 1$ 并对序列 b_n 应用练习 42。]

*44. (要求线性代数)令 A_n 是 $n \times n$ 矩阵，它的主对角线上都是 2，对角线元素旁边的所有位置是 1，其余的全是 0。找一个关于 A_n 的行列式 d_n 的递推关系。求解这个递推关系并找到一个关于 d_n 的公式。

45. 假设留在岛上的每对遗传工程培育的兔子在一个月大时生出 2 对新兔子，在两个月大和以后的每个月都生出 6 对新兔子。没有兔子死去，也没有兔子从岛上离开。

    a)一对新生的兔子留在岛上，求出与 n 个月后岛上兔子对数有关的递推关系。

    b)通过求解(a)中的递推关系确定一对新生的兔子留在岛上 n 个月以后岛上的兔子对数。

46. 假设初始在一个岛上有 2 只山羊。由于自然繁殖，岛上的山羊数每年加倍，并且每年有些山羊被带来或带走。

    a)假定每年另有 100 只山羊被放到岛上，构造一个关于第 n 年初岛上山羊数的递推关系。

    b)求解(a)的递推关系来找出第 n 年初岛上的山羊数。

    c)假定对于每个 $n \geqslant 3$，在第 n 年有 n 只山羊被从岛上带走，构造一个关于第 n 年初岛上山羊数的递推关系。

    d)求解(c)的关于第 n 年初岛上山羊数的递推关系。

47. 在一个充满活力的新软件公司，一个新女雇员的初始工资为 50 000 美元，公司允诺每年底她的工资将是她前一年工资的 2 倍，并且她在公司的每年都将额外增加 10 000 美元。

    a)构造一个与被雇用的第 n 年她的工资数有关的递推关系。

    b)求解这个递推关系找出她被雇用的第 n 年的工资。

某些线性递推关系没有常系数，但也可以被系统地求解。这就是形如 $f(n)a_n = g(n)a_{n-1} + h(n)$ 的递推关系的情况。练习 48～50 说明了这一点。

*48. a)证明递推关系

$$f(n)a_n = g(n)a_{n-1} + h(n)$$

其中 $n \geqslant 1$，$a_0 = C$，可以转变成如下形式的递推关系

$$b_n = b_{n-1} + Q(n)h(n)$$

其中 $b_n = g(n+1)Q(n+1)a_n$，满足

$$Q(n) \doteq (f(1)f(2)\cdots f(n-1))/(g(1)g(2)\cdots g(n))$$

b)使用(a)求解原来的递推关系以得到

$$a_n = \frac{C + \sum_{i=1}^{n} Q(i)h(i)}{g(n+1)Q(n+1)}$$

*49. 使用练习 48 求解递推关系 $(n+1)a_n = (n+3)a_{n-1} + n$，$n \geqslant 1$，$a_0 = 1$。

50. 可以证明当以随机的顺序排序 n 个元素时，快速排序算法(在 3.5 节练习中描述)所做的平均比较次数满足递推关系

$$C_n = n + 1 + \frac{2}{n}\sum_{k=0}^{n-1} C_k$$

$n = 1, 2, \cdots$，且初始条件 $C_0 = 0$。

a)证明 $\{C_n\}$ 也满足递推关系 $nC_n = (n+1)C_{n-1} + 2n$，$n = 1, 2, \cdots$。

b)使用练习 48 求解(a)的递推关系以找到关于 C_n 的显示公式。

**51. 证明定理 4。

**52. 证明定理 6。

53. 求解具有初始条件 $T(1)=6$ 的递推关系 $T(n)=nT^2(n/2)$。[提示：令 $n=2^k$，然后做替换 $a_k=\log T(2^k)$ 以得到一个线性非齐次的递推关系。]

6.3 分治算法和递推关系

6.3.1 引言

许多递归算法把一个给定输入的问题划分成一个或多个小问题。连续施用这种划分直到可以很快地找到这些较小问题的解。例如，在执行一个二分检索时把对一个元素在表中的搜索归约为对该元素在长度减半的表中的搜索。我们继续施用这种分解直到只剩下一个元素。当我们使用归并排序算法排序一个整数的数表时，我们将这个表划分成相等大小的两半并且分别排序每一半。然后将两个排好序的半个数表归并。这种类型的递归算法的另一个例子就是整数乘法的过程，它将两个整数相乘的问题分解成三组位数减半的整数相乘。这种分解连续施用直到只剩下一位的整数为止。这些过程叫做*分治算法*。这是因为它们将一个问题划分成较小规模的同一问题的一个或多个实例，然后用这些小问题的解来处理这个问题以找到初始问题的解，这当中也许会需要一些额外的工作。

这一节将说明怎样用递推关系分析分治算法的计算复杂性。我们将用这些递推关系估计许多不同的分治算法(包括我们在本节引入的算法)所使用的运算次数。

6.3.2 分治递推关系

假设一个递归算法把一个规模为 n 的问题分成 a 个子问题，其中每个子问题的规模是 n/b(为简单起见，假设 n 是 b 的倍数；实际上，较小的问题的规模常常是小于等于或者大于等于 n/b 的最近的整数)。此外，假设在把子问题的解组合成原来问题的解的算法处理步中需要总量为 $g(n)$ 的额外的运算。那么，如果 $f(n)$ 表示求解规模为 n 的问题所需的运算数，则得出 f 满足递推关系

$$f(n) = af(n/b) + g(n)$$

这就叫做*分治递推关系*。

首先我们将建立一些可用于研究某些重要算法的复杂性的分治递推关系。然后将说明怎样用这些分治递推关系估计这些算法的复杂性。

例 1 二分搜索 在 2.1 节我们引入了二分检索算法。当 n 是偶数时，二分检索算法把对某个元素在长度为 n 的搜索序列中的搜索转变成对同一元素在长度为 $n/2$ 的搜索序列中的二分检索。(因此，规模为 n 的问题已经被分解成规模为 $n/2$ 的问题。)为执行这个分解需要 2 次比较(一次是为了确定要用到表的哪一半，另一次是为了确定表是否还有项留下来)。所以，如果 $f(n)$是在规模为 n 的搜索序列中搜索一个元素所需要的比较次数，那么当 n 是偶数时 $f(n)=f(n/2)+2$。 ■

例 2 找一个序列的最大和最小 考虑下面的查找序列 a_1，a_2，…，a_n 中最小和最大元素的算法。如果 $n=1$，那么 a_1 就是最大和最小的元素。如果 $n>1$，把这个序列分成两个序列，或者两者有同样多的元素，或者一个集合比另一个集合多一个元素。问题就归约成查找两个较小序列的最大和最小元素。比较两个较小集合的最大和最小元素，从而得到全体的最大和最小元素，

原问题的解就得到了。

设 $f(n)$是找 n 元素集合的最小和最大元素所需要的总的比较次数。我们已经说明了当 n 是偶数时一个规模为 n 的问题可以归约成两个规模为 $n/2$ 的问题，这里要使用 2 次比较，一次是比较两个集合的最小元素，而另一次是比较两个集合的最大元素。当 n 是偶数时就得到递推关系 $f(n)=2f(n/2)+2$。 ■

例 3　归并排序　在 3.5 节介绍的归并排序算法把一个具有 n 个项(其中 n 为偶数)的待排序的表划分成两个表，每个表 $n/2$ 个元素，并且用少于 n 次的比较将两个排好序的表归并成一个排好序的表。因此，通过归并排序算法排序 n 个元素的表用到的比较次数小于 $M(n)$，其中函数 $M(n)$满足分治递推关系

$$M(n)=2M(n/2)+n$$ ■

例 4　整数的快速乘法　令人惊讶的是存在许多比整数乘法的传统算法(在 2.5 节描述过)更有效的算法。这里描述的一个有效的算法就用到了分治技术。这个快速的乘法算法开始把每个 $2n$ 位的二进制整数分成两块，每块 n 位。然后，原来 $2n$ 位的二进制整数的乘法被分解成 3 个 n 位二进制数的乘法，还要加上移位和加法。

假设 a 和 b 是两个整数的 $2n$ 位的二进制表达式(为了使得它们等长，如果需要的话，可在这些表达式前面加上若干个 0)。令

$$a=(a_{2n-1}a_{2n-2}\cdots a_1a_0)_2 \text{ 和 } b=(b_{2n-1}b_{2n-2}\cdots b_1b_0)_2$$

令

$$a=2^nA_1+A_0,\quad b=2^nB_1+B_0$$

其中

$$A_1=(a_{2n-1}\cdots a_{n+1}a_n)_2,\qquad A_0=(a_{n-1}\cdots a_1a_0)_2$$
$$B_1=(b_{2n-1}\cdots b_{n+1}b_n)_2,\qquad B_0=(b_{n-1}\cdots b_1b_0)_2$$

快速整数乘法算法是基于恒等式

$$ab=(2^{2n}+2^n)A_1B_1+2^n(A_1-A_0)(B_0-B_1)+(2^n+1)A_0B_0$$

关于这个恒等式的一个重要的事实就是，它证明了两个 $2n$ 位整数的乘法可以用 3 个 n 位整数的乘法加上加法、减法以及移位来实现。这证明了如果 $f(n)$是两个 n 位整数相乘所需的按位运算的总数，那么

$$f(2n)=3f(n)+Cn$$

这个等式成立的理由如下：3 次 n 位整数的乘法可以使用 $3f(n)$次按位运算实现。每次加法、减法和移位使用的运算次数是 n 位运算的常数倍，而 Cn 表示由这些运算用到的总的按位运算数。 ■

例 5　快速矩阵乘法　在 2.7 节例 5 中我们证明了使用矩阵乘法的定义进行两个 $n\times n$ 矩阵相乘需要 n^3 次乘法和 $n^2(n-1)$次加法。因此，按照这种方法计算两个 $n\times n$ 矩阵之积需要 $O(n^3)$次运算(乘法和加法)。令人惊讶的是，对于两个 $n\times n$ 矩阵相乘存在更有效的分治算法。这个由 V. 斯特拉森于 1969 年提出的算法当 n 是偶数时将两个 $n\times n$ 矩阵的相乘归约为两个$(n/2)\times(n/2)$矩阵的 7 次相乘和$(n/2)\times(n/2)$矩阵的 15 次相加。(要了解这个算法的细节见[CoLeRiSt01]。)于是，如果 $f(n)$是用到的运算(乘法和加法)次数，那么当 n 是偶数时有

$$f(n)=7f(n/2)+15n^2/4$$ ■

正如例 1～例 5 所示，在许多不同的情况中都出现了形如 $f(n)=af(n/b)+g(n)$的递推关系。可以对满足这种递推关系的函数的阶作出估计。假设 f 满足这个递推关系，其中 n 可被 b 整除。令 $n=b^k$，其中 k 是一个正整数。那么

$$f(n)=af(n/b)+g(n)$$

$$= a^2 f(n/b^2) + ag(n/b) + g(n)$$
$$= a^3 f(n/b^3) + a^2 g(n/b^2) + ag(n/b) + g(n)$$
$$\vdots$$
$$= a^k f(n/b^k) + \sum_{j=0}^{k-1} a^j g(n/b^j)$$

由于 $n/b^k=1$，从而有

$$f(n) = a^k f(1) + \sum_{j=0}^{k-1} a^j g(n/b^j)$$

我们可以使用这个关于 $f(n)$ 的等式估计满足分治关系的函数的阶。

定理 1 设 f 是满足递推关系

$$f(n) = af(n/b) + c$$

的增函数，其中 n 被 b 整除，$a\geqslant 1$，b 是大于 1 的整数，c 是一个正实数。那么

$$f(n) \text{ 是} \begin{cases} O(n^{\log_b a}) & \text{如果 } a > 1 \\ O(\log n) & \text{如果 } a = 1 \end{cases}$$

进一步，当 $n=b^k$，其中 k 是正整数，

$$f(n) = C_1 n^{\log_b a} + C_2$$

其中 $C_1=f(1)+c/(a-1)$ 且 $C_2=-c/(a-1)$。

证 首先令 $n=b^k$。由定理前面的讨论得到的关于 $f(n)$ 的表达式和 $g(n)=c$，我们有

$$f(n) = a^k f(1) + \sum_{j=0}^{k-1} a^j c = a^k f(1) + c\sum_{j=0}^{k-1} a^j$$

先考虑 $a=1$ 的情况。那么有

$$f(n) = f(1) + ck$$

由于 $n=b^k$，我们有 $k=\log_b n$，于是

$$f(n) = f(1) + c\log_b n$$

当 n 不是 b 的幂时，对某个正整数 k 有 $b^k<n<b^{k+1}$。由于 f 是增加的，故 $f(n)\leqslant f(b^{k+1})=f(1)+c(k+1)=(f(1)+c)+ck\leqslant(f(1)+c)+c\log_b n$。因此，当 $a=1$ 时在两种情况下 $f(n)$ 都是 $O(\log n)$。

现在假设 $a>1$。首先假定 $n=b^k$，k 是正整数。由几何级数的求和公式(3.2 节定理 1)得

$$\begin{aligned} f(n) &= a^k f(1) + c(a^k-1)/(a-1) \\ &= a^k[f(1)+c/(a-1)] - c/(a-1) \\ &= C_1 n^{\log_b a} + C_2 \end{aligned}$$

这是由于 $a^k=a^{\log_b n}=n^{\log_b a}$(见附录 1 的练习 4)，其中 $C_1=f(1)+c/(a-1)$ 且 $C_2=-c/(a-1)$。

现在假设 n 不是 b 的幂。那么 $b^k<n<b^{k+1}$，其中 k 是一个非负整数。由于 f 是增加的，

$$\begin{aligned} f(n) &\leqslant f(b^{k+1}) = C_1 a^{k+1} + C_2 \\ &\leqslant (C_1 a) a^{\log_b n} + C_2 \\ &\leqslant (C_1 a) n^{\log_b a} + C_2 \end{aligned}$$

这是由于 $k\leqslant\log_b n<k+1$。

于是，$f(n)$ 是 $O(n^{\log_b a})$。 □

例 6～例 9 说明怎样使用定理 1。

例 6 设 $f(n)=5f(n/2)+3$ 且 $f(1)=7$，求 $f(2^k)$，其中 k 是一个正整数。如果 f 是一个增函数，请估计 $f(n)$。

解 从定理 1 的证明，考虑 $a=5$，$b=2$，$c=3$，我们看到如果 $n=2^k$，那么

$$\begin{aligned} f(n) &= a^k[f(1)+c/(a-1)]+[-c/(a-1)] \\ &= 5^k[7+(3/4)]-3/4 \\ &= 5^k(31/4)-3/4 \end{aligned}$$

又如果 $f(n)$是增加的，定理 1 证明了 $f(n)$是 $O(n^{\log_b a})=O(n^{\log 5})$。 ■

我们可以使用定理 1 估计二分检索算法和例 2 查找序列的最小和最大元素的算法的计算复杂性。

例 7 估计二分检索使用的比较次数。

解 在例 1 中证明了当 n 是偶数时 $f(n)=f(n/2)+2$，其中 $f(n)$是在规模为 n 的序列中实现一个二分检索需要的比较次数。因此得出 $f(n)$是 $O(\log n)$。 ■

例 8 估计用例 2 给定的算法查找序列的最大和最小元素所使用的比较次数。

解 在例 2 我们证明了当 n 是偶数时 $f(n)=2f(n/2)+2$，其中 f 是算法需要的比较次数。于是，由定理 1 得到 $f(n)=O(n^{\log 2})=O(n)$。 ■

我们现在叙述一个更一般的、更复杂的定理，定理 1 是它的特例。这个定理(或者更强的版本)有时称作主定理(MasterTheorem)，因为它在分析许多重要的分治算法的复杂性中很有用。

定理 2 主定理 设 f 是满足递推关系

$$f(n) = af(n/b) + cn^d$$

的增函数，其中 $n=b^k$，k 是一个正整数，$a\geqslant 1$，b 是大于 1 的整数，c 和 d 是实数，满足 c 是正的且 b 是非负的。那么

$$f(n) \text{ 是} \begin{cases} O(n^d) & \text{如果 } a < b^d \\ O(n^d \log n) & \text{如果 } a = b^d \\ O(n^{\log_b a}) & \text{如果 } a > b^d \end{cases}$$

定理 2 的证明留给读者作为本节末的练习 29～33。

例 9 **归并排序的复杂性** 在例 3 我们解释了用归并排序来排序 n 个元素的表使用的比较次数少于 $M(n)$，其中 $M(n)=2M(n/2)+n$。根据主定理(定理 2)我们找到 $M(n)$是 $O(n\log n)$，这与在 3.5 节得到的估计一致。 ■

例 10 估计使用例 4 描述的快速乘法算法做两个 n 位整数相乘所需要的按位运算次数。

解 例 4 证明了当 n 是偶数时 $f(n)=3f(n/2)+Cn$，其中 $f(n)$是使用快速乘法算法做两个 n 位整数相乘所需的按位运算次数。于是，由定理 2 得到 $f(n)$是 $O(n^{\log 3})$。注意 $\log 3\approx 1.6$。因为通常的乘法算法使用 $O(n^2)$次按位运算，所以对于足够大的整数，包括实际应用中出现的大整数，快速乘法算法在时间复杂性方面比通常的算法有了本质的改进。 ■

例 11 估计使用例 5 的矩阵乘法算法做两个 $n\times n$ 矩阵相乘所需要的乘法和加法次数。

解 令 $f(n)$表示使用例 5 提到的算法做两个 $n\times n$ 矩阵相乘所需的加法和乘法次数。当 n 是偶数时我们有 $f(n)=7f(n/2)+15n^2/4$。于是由定理 2 得到 $f(n)$是 $O(n^{\log 7})$。注意 $\log 7\approx 2.8$。由于通常的两个 $n\times n$ 矩阵相乘的算法要用 $O(n^3)$次加法和乘法，显然，对足够大的整数 n，包括出现在许多实际应用中的大整数，这个算法比通常的算法在时间复杂性方面更加有效。 ■

最接近点对问题 我们在结束这一节之前引入一个来自于计算几何的分治算法，计算几何是离散数学的一部分，专注于求解几何问题的算法。

例 12 最接近点对问题 平面上有 n 个点(x_1, y_1), (x_2, y_2), …, (x_n, y_n)的集合,其中两点(x_i, y_i)和(x_j, y_j)之间的距离是通常的欧几里得距离$\sqrt{(x_i-x_j)^2+(y_i-y_j)^2}$。考虑确定其中最接近的点对的问题。这个问题见于许多应用中,例如确定在某航空控制中心管理的特定高度的空间最近的一对飞机。怎样以一种有效的方式找到这个最近的点对?

解 为解决这个问题,可以首先确定每对点的距离,然后找到这些距离的最小值。但是,这种方法需要 $O(n^2)$次的距离计算和比较,因为存在 $C(n, 2)=n(n-1)/2$ 个点对。不过存在一个精致的分治算法,对于 n 个点可以用 $O(n\log n)$次的距离计算和比较求解这个最接近点对问题。我们在这里描述的算法是归功于米凯尔·萨莫斯(见[PrSa85])。

为了简单起见,假设 $n=2^k$,其中 k 是正整数。(我们避免某些当 n 不是 2 的幂时必需的技术上的考虑)。当 $n=2$ 时,只有一对点;在这两个点之间的距离就是最小距离。

算法的递归部分将问题划分成两个子问题,每个涉及一半的点。第一步使用归并排序算法,根据这些点的 x 坐标对点进行排序。这需要 $O(n\log n)$次运算。一旦根据它们的 x 坐标对这些点排好序,我们画一条垂线 ℓ 将 n 个点分成两部分,左半部分和右半部分大小相等,每部分包含 $n/2$ 个点,如图 6-5 所示。(如果有任何点落到划分线 ℓ 上,必要的话,我们把它们分在这两部分里。)在后面的递归步骤我们不再需要根据 x 坐标排序,因为我们可以从所有的点中选择对应的排序子集。这个选择是可以用 $O(n)$次比较完成的任务。

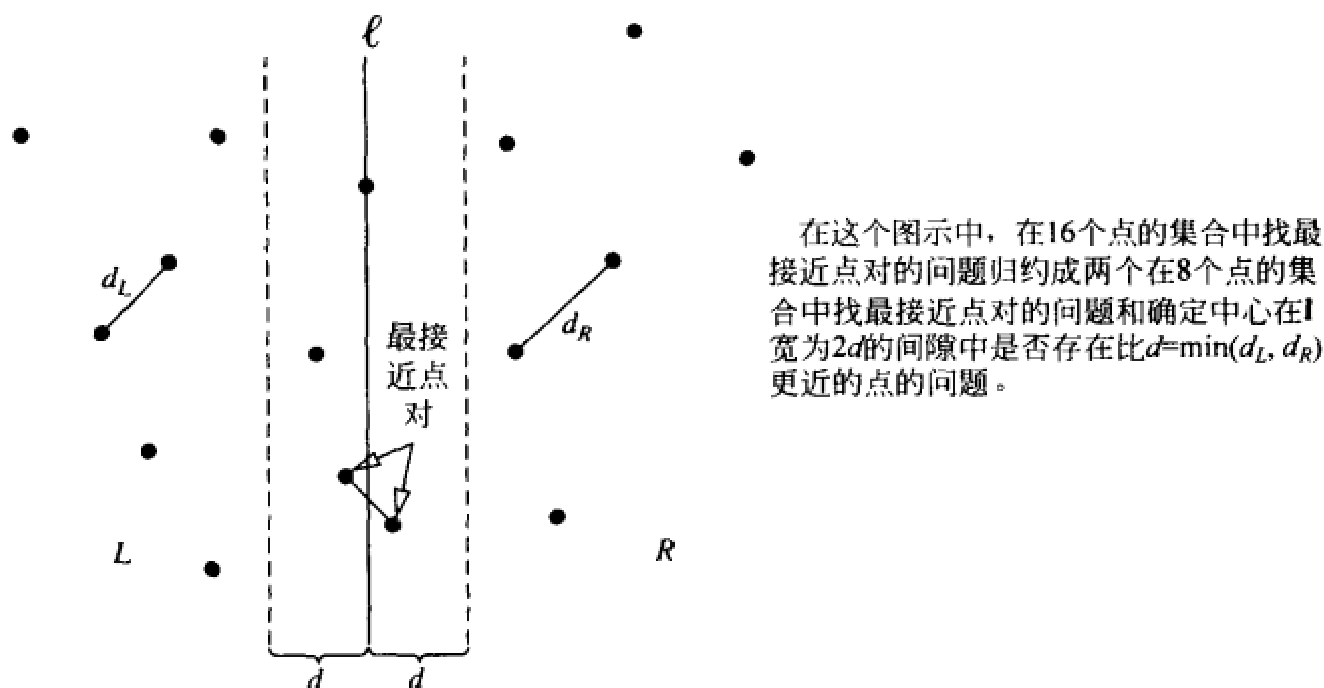


图 6-5 求解最接近点对问题的算法的递归步

最接近的点对的位置有三种可能:(1)它们两点都在左部区域 L,(2)它们两点都在右部区域 R,(3)一点在左部区域且另一点在右部区域。递归地使用这个算法计算 d_L 和 d_R,其中 d_L 是在左部区域的点之间的最小距离,d_R 是在右部区域的点之间的最小距离,令 $d=\min(d_L, d_R)$。为了成功地将在原始集合找最接近点对的问题划分成在两个区域分别找最短距离的问题,我们必须处理算法的治之部分,这要求我们考虑最接近的点处在不同的区域的情况,即一点在 L 中,另一点在 R 中。因为存在一对点距离为 d,或者它们都在 R 中,或者它们都在 L 中。对于分在不同区域的最接近的点,要求其距离一定小于 d。

如果一点在左边区域,一点在右边区域且处在小于 d 的距离内,这些点一定位于宽度 $2d$ 的以 l 作为其中心的垂直间隙中。(否则,这些点的距离一定大于它们的 x 坐标之差,而这个距离将超过 d。)为了检查在这个间隙中的点,我们对它们进行排序并按照 y 坐标递增的顺序把它们列出来。这可以使用归并排序用 $O(n\log n)$次运算做到,并且只需在算法开始时做一次,而不是在

每个递归步做。在每个递归步，我们从已经按照其 y 坐标排好序的所有点的集合，构造在这个区域内的根据其 y 坐标排序的点的子集，这可以用 $O(n)$ 次比较做到。

从间隙中具有最小 y 坐标的一个点开始，我们连续地检查间隙中的每个点，计算这个点与间隙中所有其他具有较大 y 坐标且与这个点的距离小于 d 的点之间的距离。注意到为检查点 p，我们只需要考虑在 p 和下述矩形中的一组点之间的距离。这个矩形的高为 d，宽为 $2d$，p 在它的底边上，并且它的垂直边与 ℓ 的距离为 d。

我们可以证明在这个点集中至多存在 8 个点，其中包含 p 在内(或者在这个 $2d\times d$ 的矩形的边上)。为了看到这一点，注意到在图 6-6 所示的 8 个 $d/2\times d/2$ 的正方形中，每个正方形内部至多可能存在一个点。这是由于在一个正方形的边上或者内部最远距离的点是对角线的长度 $d/\sqrt{2}$ (使用勾股定理可以得到)，这个距离小于 d，并且每个这样的正方形是完全处在左区域或者右区域内。这意味着在这一步我们至多只需要与 d 比较 7 个距离，这些距离是在 p 和矩形内部或者边上 7 个或者更少的其他的点之间。

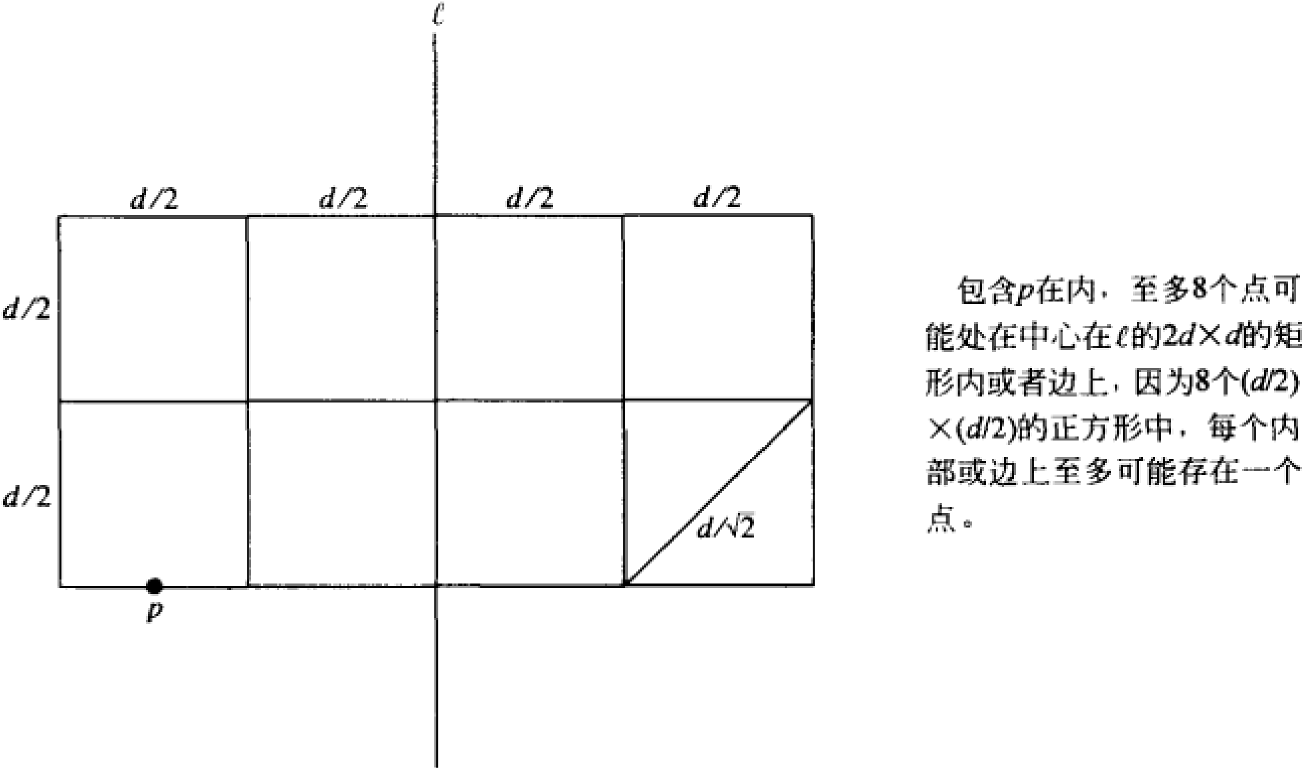


图 6-6　说明对间隙中的每个点至多需要考虑另外 7 个点

由于在宽为 $2d$ 的间隙中的总点数不超过 n(集合中的总点数)，至多需要与 d 比较 $7n$ 个距离以找到点之间的最小距离。即只存在 $7n$ 个距离可能小于 d。因此，一旦用归并排序按照这些点的 x 坐标和 y 坐标对它们进行排序以后，我们发现求解最接近点对问题需要的比较次数不超过满足递推关系

$$f(n) = 2f(n/2) + 7n$$

的增函数 $f(n)$，其中 $f(2)=1$。根据定理 2，得到 $f(n)$是 $O(n\log n)$。用归并排序算法根据点的 x 坐标和 y 坐标对点做两次排序，做每次排序用 $O(n\log n)$次比较，在算法的 $O(\log n)$个步中的每一步，这些坐标的排序子集每次可以用 $O(n)$ 次比较得到。因此，这个最接近点对问题可以用 $O(n\log n)$次比较求解。■

练习

1. 在 64 个元素的集合中，做二分检索需要多少次比较？
2. 在 128 个元素的序列中，使用例 2 中的算法查找最大和最小的元素需要多少次比较？

3. 使用快速乘法算法乘$(1110)_2$和$(1010)_2$。
4. 用伪码表示快速乘法算法。
5. 确定在例4中的常数C的值，并且使用它估计用快速乘法算法做两个64位二进制整数相乘所需要的按位运算次数。
6. 用例4引入的算法做两个32×32矩阵相乘需要多少次运算？
7. 假设当n被3整除时有$f(n)=f(n/3)+1$和$f(1)=1$，求
   a) $f(3)$ b) $f(27)$ c) $f(729)$
8. 假设当n是偶数时有$f(n)=2f(n/2)+3$和$f(1)=5$，求
   a) $f(2)$ b) $f(8)$ c) $f(64)$ d) $f(1024)$
9. 假设当n被5整除时有$f(n)=f(n/5)+3n^2$和$f(1)=4$，求
   a) $f(5)$ b) $f(125)$ c) $f(3125)$
10. 当$n=2^k$时求$f(n)$，其中f满足递推关系$f(n)=f(n/2)+1$，$f(1)=1$。
11. 如果f是一个增函数，估计练习10中的f的规模。
12. 当$n=3^k$时求$f(n)$，其中f满足递推关系$f(n)=2f(n/3)+4$，$f(1)=1$。
13. 如果f是一个增函数，估计练习12中的f的规模。
14. 假设在一个淘汰锦标赛中有$n=2^k$个队，其中在第一轮有$n/2$场比赛，$n/2=2^{k-1}$个赢的队进入第二轮比赛，依此进行。建立一个关于锦标赛的轮数的递推关系。
15. 在练习14的淘汰锦标赛中如果有32个队，需要进行多少轮比赛？
16. 求解练习14所描述的关于锦标赛轮数的递推关系。
17. 假设n个投票人为不同的候选人(可能存在多于2个候选人)进入某个办公室投票，选票作为一个序列的元素。如果一个人得到的选票超过半数他就赢得竞选。
    a) 设计一个分治算法确定是否一个候选人得到半数选票，如果是，则确定这个候选人是谁。[提示：设n为偶数，并且将选票序列划分成两个序列，每个序列具有$n/2$个元素。注意到如果对于两个半长序列每一个都没有得到一半以上的选票，那么一个人就不可能得到所有选票的一半以上。]
    b) 使用主定理估计你在(a)中设计的算法所需要的比较次数。
18. 假设在一组n个人中，每个人从候选人的提名中恰好选两个人担任一个委员会的两个职务。只要每人得到超过$n/2$的选票，这头两个人将赢得这两个席位。
    a) 设计一个分治算法，确定两个得到最多选票的候选人是否每个人至少得到了$n/2$的选票。如果是，确定这两个候选人是谁。
    b) 使用主定理估计你在(a)中设计的算法所需要的比较次数。
19. a) 使用3.5节中的练习18的递归算法为计算x^n所需要的乘法次数建立一个分治递推关系，其中x为实数，n是正整数。
    b) 使用你在(a)中找到的递推关系构造使用递归算法计算x^n所用乘法次数的大O估计。
20. a) 使用3.5节中的例2的递归算法为计算$a^n \bmod m$所需要的模乘法次数建立一个分治递推关系，其中a，m，n为正整数。
    b) 使用你在(a)中找到的递推关系构造使用递归算法计算$a^n \bmod m$所用模乘法次数的大O估计。
21. 设函数f满足递推关系$f(n)=2f(\sqrt{n})+1$，其中n是大于1的完全平方数且$f(2)=1$。
    a) 求$f(16)$。
    b) 求关于$f(n)$的大O估计。[提示：做替换$m=\log n$。]
22. 设函数f满足递推关系$f(n)=2f(\sqrt{n})+\log n$，其中n是大于1的完全平方数且$f(2)=1$。
    a) 求$f(16)$。
    b) 求关于$f(n)$的大O估计。[提示：做替换$m=\log n$。]

**23. 这个练习涉及求n个实数序列的连续项的最大和问题。当所有的项都是正数时，所有项之和就给出了答案，但是当某些项是负数时情况就比较复杂了。例如，序列-2，3，-1，6，-7，4的连续项的最

大和是 3+(−1)+6=8。(这个练习基于[Be86])。

a)使用伪码描述一个求解该问题的算法，这个算法依次找从第一项开始的连续项之和，从第二项开始的连续项之和，等等，并在算法执行时记录当前找到的最大和。

b)依照所做的计算和的次数与比较次数确定在(a)中算法的计算复杂性。

c)设计一个分治算法求解这个问题。[提示：假设序列中有偶数个项，把这个序列分成两半。解释当连续项的最大和包含了在两个半序列的项时怎样处理这种情况。]

d)使用(c)中的算法求下面每个序列的连续项的最大和：−2，4，−1，3，5，−6，1，2；4，1，−3，7，−1，−5，3，−2；−1，6，3，−4，−5，8，−1，7。

e)对于(c)中的分治算法使用的求和次数和比较次数求一个递推关系。

f)使用主定理估计这个分治算法的计算复杂性。依照计算复杂性把这个算法和(a)中的算法做比较，结果如何？

24. 使用例 12 描述的求最接近点对的算法，使用点之间的欧几里得距离，求下述点(1，3)，(1，7)，(2，4)，(2，9)，(3，1)，(3，5)，(4，3)，(4，7)的最接近点对。

25. 使用例 12 描述的求最接近点对的算法，使用点之间的欧几里得距离，求下述点(1，2)，(1，6)，(2，4)，(2，8)，(3，1)，(3，6)，(3，10)，(4，3)，(5，1)，(5，5)，(5，9)，(6，7)，(7，1)，(7，4)，(7，9)，(8，6)的最接近点对。

*26. 使用伪码描述例 12 中所叙述的解最接近点对问题的递归算法。

27. 如果两点间的距离定义是 $d((x_i, y_i), (x_j, y_j)) = \max(|x_i - x_j|, |y_i - y_j|)$，使用例 12 描述的算法中用的那些合理的步骤并加以改变，构造一个求两点之间的最小距离的算法。

*28. 设一个人从 n 个数的集合中取一个数 x，第二个人通过连续选取的 n 个数的子集猜这个数。他问第一个人是否 x 在每个集合里。第一个人回答“是”或者“不是”。当第一个人每次回答都是真话时，我们可以使用 $\log n$ 次询问找到 x。做法是连续在每次询问时将这个集合对半划分。1976 年由斯坦尼斯劳·乌拉姆(Stanislaw Ulam)提出的乌拉姆问题是：假设允许第一个人恰好说谎一次，为找到 x 需要多少次询问？

a)给定数 x 和 n 个元素的集合，证明每个问题问 2 次并且当发现说谎时可以多问一个问题，那么乌拉姆问题可以用 $2\log n+1$ 次询问求解。

b)把初始的 n 元素集合划分成 4 部分，每部分具有 $n/4$ 个元素，证明使用 2 次询问就可以排除 1/4 的元素。[提示：使用 2 次询问，其中每次询问都问是否这个元素在两个 $n/4$ 个元素的子集的并集之中，并且其中一个 $n/4$ 个元素的子集出现在两次询问中。]

c)利用(b)证明如果 $f(n)$ 等于用(b)中的方法求解乌拉姆问题所用到的询问次数，且 n 被 4 整除，那么 $f(n)=f(3n/4)+2$。

d)求解(c)中关于 $f(n)$ 的递推关系。

e)这种每个问题问两次来求解乌拉姆问题的天真的方法与基于(b)的分治方法，哪种具有更高的效率？求解乌拉姆问题的最有效的方法已经由 A. 派尔克(Pelc)确定[Pe87]。

在练习 29～33 中，假设 f 是一个满足递推关系 $f(n)=af(n/b)+cn^d$ 的增函数，$a\geq 1$，b 是大于 1 的整数，c 和 d 是正实数。这些练习提供一个关于定理 2 的证明。

*29. 证明如果 $a=b^d$ 且 n 是 b 的幂，那么 $f(n)=f(1)n^d+cn^d\log_b n$。

30. 使用练习 29 证明如果 $a=b^d$，那么 $f(n)$ 是 $O(n^d\log n)$。

*31. 证明如果 $a\neq b^d$ 且 n 是 b 的幂，那么 $f(n)=C_1n^d+C_2n^{\log_b a}$，其中 $C_1=b^dc/(b^d-a)$ 且 $C_2=f(1)+b^dc/(a-b^d)$。

32. 使用练习 31 证明如果 $a<b^d$，那么 $f(n)$ 是 $O(n^d)$。

33. 使用练习 31 证明如果 $a>b^d$，那么 $f(n)$ 是 $O(n^{\log_b a})$。

34. 当 $n=4^k$ 时，求 $f(n)$，其中 f 满足递推关系 $f(n)=5f(n/4)+6n$，$f(1)=1$。

35. 如果 f 是增函数，估计练习 34 中 f 的大小。
36. 当 $n=2^k$ 时，求 $f(n)$，其中 f 满足递推关系 $f(n)=8f(n/2)+n^2$，$f(1)=1$。
37. 如果 f 是增函数，估计练习 36 中 f 的大小。

6.4 生成函数

6.4.1 引言

表示序列的一种有效方法就是生成函数，它把序列的项作为一个形式幂级数中变量 x 的幂的系数。可以用生成函数求解许多类型的计数问题，例如在各种限制下选取或分配不同种类物体的方式数，使用不同面额的硬币换一美元的方式数等。也可以用生成函数求解递推关系。它先把关于序列的项的递推关系转换成涉及生成函数的方程，然后求解这个方程并找出关于这个生成函数的封闭形式。从这个封闭形式可以找到生成函数的幂级数的系数，从而求解了原来的递推关系。生成函数也可以利用函数之间相对简单的关系来证明组合恒等式，由于这些关系可以转换成涉及序列的项的恒等式。生成函数是有用的工具，除了本节描述的内容以外，可以用它来研究序列的许多性质，例如建立关于序列的项的渐进公式。

我们从序列的生成函数的定义开始。

定义 1 实数序列 a_0，a_1，…，a_k，…的生成函数是无穷级数

$$G(x) = a_0 + a_1x + \cdots + a_kx^k + \cdots = \sum_{k=0}^{\infty} a_kx^k$$

注意 定义 1 给出的 $\{a_k\}$ 的生成函数有时叫做 $\{a_k\}$ 的普通的生成函数，以和这个序列的其他类型的生成函数相区别。

例 1 序列 $\{a_k\}$ 具有 $a_k=3$，$a_k=k+1$ 和 $a_k=2^k$ 的生成函数分别是 $\sum_{k=0}^{\infty} 3x^k$，$\sum_{k=0}^{\infty}(k+1)x^k$，$\sum_{k=0}^{\infty} 2^kx^k$。■

我们通过置 $a_{n+1}=0$，$a_{n+2}=0$，依此下去，把一个有限序列 a_0，a_1，…，a_n 扩充成一个无限序列，就可以定义一个实数的有限序列的生成函数。这个无限序列 $\{a_n\}$ 的生成函数 $G(x)$ 是一个 n 次多项式，因为当 $j>n$ 时没有形如 a_jx^j 的项出现，即

$$G(x) = a_0 + a_1x + \cdots + a_nx^n$$

例 2 序列 1，1，1，1，1，1 的生成函数是什么？

解 1，1，1，1，1，1 的生成函数是

$$1 + x + x^2 + x^3 + x^4 + x^5$$

由 3.2 节的定理 1 有

$$(x^6 - 1)/(x - 1) = 1 + x + x^2 + x^3 + x^4 + x^5$$

因此，$G(x)=(x^6-1)/(x-1)$ 是序列 1，1，1，1，1，1 的生成函数。■

例 3 设 m 是正整数。令 $a_k=C(m, k)$，$k=0$，1，2，…，m。那么序列 a_0，a_1，…，a_m 的生成函数是什么？

解 这个序列的生成函数是

$$G(x) = C(m,0) + C(m,1)x + C(m,2)x^2 + \cdots + C(m,m)x^m$$ ■

二项式定理证明 $G(x)=(1+x)^m$。

6.4.2 关于幂级数的有用事实

当用生成函数求解计数问题时，通常将它们考虑成形式幂级数。这里忽略了这些级数的收敛问题。但是为了应用某些微积分的结果，考虑幂级数对什么 x 收敛有时是很重要的。在我们的讨论中将不涉及收敛性问题。熟悉微积分的读者如果想了解所涉及级数的收敛性的细节，可以参阅有关这方面内容的教科书。

现在我们将叙述某些与无穷级数有关的重要事实，这些将在研究生成函数时用到。这些事实的讨论和相关的结果都可以在微积分教科书中找到。

例 4 函数 $f(x)=1/(1-x)$ 是序列 1，1，1，1，…的生成函数，因为对 $|x|<1$ 有

$$1/(1-x) = 1 + x + x^2 + \cdots$$

■

例 5 函数 $f(x)=1/(1-ax)$ 是序列 1，a，a^2，a^3，…的生成函数，因为当 $|ax|<1$ 或等价地说对 $|x|<1/|a|$，$a\neq 0$ 有

$$1/(1-ax) = 1 + ax + a^2x^2 + \cdots$$

■

我们也需要了解两个生成函数是怎样相加和相乘的。这些结果的证明也可以在微积分教科书中找到。

定理 1 令 $f(x) = \sum_{k=0}^{\infty} a_k x^k$，$g(x) = \sum_{k=0}^{\infty} b_k x^k$，那么

$$f(x) + g(x) = \sum_{k=0}^{\infty}(a_k + b_k)x^k, \quad f(x)g(x) = \sum_{k=0}^{\infty}\left(\sum_{j=0}^{k} a_j b_{k-j}\right)x^k$$

我们将在例 6 中说明怎样使用定理 1。

例 6 设 $f(x)=1/(1-x)^2$。用例 4 求出表达式 $f(x) = \sum_{k=0}^{\infty} a_k x^k$ 中的系数 a_0，a_1，a_2，…。

解 由例 4 看出

$$1/(1-x) = 1 + x + x^2 + x^3 + \cdots$$

因此，由定理 1 有

$$1/(1-x)^2 = \sum_{k=0}^{\infty}\left(\sum_{j=0}^{k} 1\right)x^k = \sum_{k=0}^{\infty}(k+1)x^k$$

■

注意 这一结果也可以通过微分从例 4 导出。从已知生成函数的恒等式产生新的恒等式的一种有用的技术就是求导。

为了用生成函数求解许多重要的计数问题，我们需要在指数不是正整数的情况下应用二项式定理。在叙述广义二项式定理之前，我们需要定义广义二项式系数。

定义 2 设 u 是实数且 k 是非负整数，那么广义二项式系数 $\binom{u}{k}$ 定义为

$$\binom{u}{k} = \begin{cases} u(u-1)\cdots(u-k+1)/k! & \text{若 } k > 0 \\ 1 & \text{若 } k = 0 \end{cases}$$

例 7 求广义二项式系数 $\binom{-2}{3}$ 和 $\binom{1/2}{3}$ 的值。

解 在定义 2 中取 $u=-2$ 和 $k=3$ 得

$$\binom{-2}{3} = \frac{(-2)(-3)(-4)}{3!} = -4$$

类似地取 $u=1/2$ 和 $k=3$ 得

$$\binom{1/2}{3}=\frac{(1/2)(1/2-1)(1/2-2)}{3!}$$
$$=(1/2)(-1/2)(-3/2)/6$$
$$=1/16$$ ■

当上边的参数是负整数时，例 8 对广义二项式系数提供了一个有用的公式。我们后面的讨论中会用到它。

例 8 当上边的参数是负整数时，广义二项式系数可以用通常的二项式系数的项表示。为此只需注意到

$$\binom{-n}{r}=\frac{(-n)(-n-1)\cdots(-n-r+1)}{r!}$$
$$=\frac{(-1)^r n(n+1)\cdots(n+r-1)}{r!}$$
$$=\frac{(-1)^r(n+r-1)(n+r-2)\cdots n}{r!}$$
$$=\frac{(-1)^r(n+r-1)!}{r!(n-1)!}$$
$$=(-1)^r\binom{n+r-1}{r}$$
$$=(-1)^r C(n+r-1,r)$$ ■

我们现在叙述广义二项式定理。

定理 2 广义二项式定理 设 x 是实数，$|x|<1$，u 是实数，那么

$$(1+x)^u=\sum_{k=0}^{\infty}\binom{u}{k}x^k$$

可以使用麦克劳林级数的理论证明定理 2，我们将这个证明留给熟悉这部分微积分的读者完成。

注意 当 u 是正整数时，广义二项式定理就归约到 4.4 节提出的二项式定理，因为如果 $k>u$，在这种情况下 $\binom{u}{k}=0$。

例 9 说明了当指数是负整数时定理 2 的应用。

例 9 当 n 是正整数时，使用广义二项式定理求 $(1+x)^{-n}$ 和 $(1-x)^{-n}$ 的生成函数。

解 由广义二项式定理得

$$(1+x)^{-n}=\sum_{k=0}^{\infty}\binom{-n}{k}x^k$$

使用例 8 所提供的关于 $\binom{-n}{k}$ 的简单公式得到

$$(1+x)^{-n}=\sum_{k=0}^{\infty}(-1)^k C(n+k-1,k)x^k$$

用 $-x$ 代替 x 得到

$$(1-x)^{-n}=\sum_{k=0}^{\infty}C(n+k-1,k)x^k$$ ■

表 6-1 归纳了一些经常出现的有用的生成函数。

表 6-1 有用的生成函数

$G(x)$	a_k
$(1+x)^n = \sum_{k=0}^{n} C(n,k)x^k = 1 + C(n,1)x + C(n,2)x^n + \cdots + x^n$	$C(n,k)$
$(1+ax)^n = \sum_{k=0}^{n} C(n,k)a^k x^k = 1 + C(n,1)ax + C(n,2)a^2x^2 + \cdots + a^n x^n$	$C(n,k)a^k$
$(1+x^r)^n = \sum_{k=0}^{n} C(n,k)x^{rk} = 1 + C(n,1)x^r + C(n,2)x^{2r} + \cdots + x^{rn}$	当 r/k 时为 $C(n,k/r)$，否则为 0
$\frac{1-x^{n+1}}{1-x} = \sum_{k=0}^{n} x^k = 1 + x + x^2 + \cdots + x^n$	当 $k \leqslant n$ 时为 1，否则为 0
$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + \cdots$	1
$\frac{1}{1-ax} = \sum_{k=0}^{\infty} a^k x^k = 1 + ax + a^2x^2 + \cdots$	a^k
$\frac{1}{1-x^r} = \sum_{k=0}^{\infty} x^{rk} = 1 + x^r + x^{2r} + \cdots$	当 r/k 时为 1，否则为 0
$\frac{1}{(1-x)^2} = \sum_{k=0}^{\infty} (k+1)x^k = 1 + 2x + 3x^2 + \cdots$	$k+1$
$\frac{1}{(1-x)^n} = \sum_{k=0}^{\infty} C(n+k-1,k)x^k = 1 + C(n,1)x + C(n+1,2)x^2 + \cdots$	$C(n+k-1,k) = C(n+k-1,n-1)$
$\frac{1}{(1+x)^n} = \sum_{k=0}^{\infty} C(n+k-1,k)(-1)^k x^k = 1 - C(n,1)x + C(n+1,2)x^2 - \cdots$	$(-1)^k C(n+k-1,k) = (-1)^k C(n+k-1,n-1)$
$\frac{1}{(1-ax)^n} = \sum_{k=0}^{\infty} C(n+k-1,k)a^k x^k = 1 + C(n,1)ax + C(n+1,2)a^2x^2 + \cdots$	$C(n+k-1,k)a^k = C(n+k-1,n-1)a^k$
$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$	$1/k!$
$\ln(1+x) = \sum_{k=0}^{\infty} \frac{(-1)^{k+1}}{k} x^k = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$	$(-1)^{k+1}/k$

注：在大多数微积分的书中讨论幂级数的部分可以找到关于最后两个生成函数的级数。

6.4.3 计数问题与生成函数

生成函数可以用于求解各种计数问题。特别地，它们可以用于计数各种类型的组合数。在第 4 章，我们开发了一些技术计数 n 元素集合的允许重复的 r 组合数，在这些组合中可能存在某些附加的约束。这种问题与计数形如

$$e_1 + e_2 + \cdots + e_n = C$$

方程的解是等价的，其中 C 是常数，每个 e_i 是可能具有某些约束的非负整数。也可以用生成函数求解这种类型的计数问题，正如例 10～例 12 所示。

例 10 求

$$e_1 + e_2 + e_3 = 17$$

的解的个数，其中 e_1，e_2，e_3 是非负整数，满足 $2 \leqslant e_1 \leqslant 5$，$3 \leqslant e_2 \leqslant 6$，$4 \leqslant e_3 \leqslant 7$。

解 具有上述限制的解的个数是

$$(x^2 + x^3 + x^4 + x^5)(x^3 + x^4 + x^5 + x^6)(x^4 + x^5 + x^6 + x^7)$$

的展开式中 x^{17} 的系数。这是因为我们在乘积中得到等于 x^{17} 的项是通过在第一个和中取项 x^{e_1}，在第二个和中取项 x^{e_2}，在第三个和中取项 x^{e_3}，其中幂指数 e_1，e_2 和 e_3 满足方程 $e_1 + e_2 + e_3 = 17$ 和给定的限制。

不难看出在这个乘积中的 x^{17} 的系数是 3。因此，存在 3 个解。（注意，计算这个系数与枚举方程的具有给定约束的所有解几乎要做同样多的工作。但是，正如我们将看到的，这里说明的方法常常可以用于求解各种各样的具有特殊规则的计数问题。此外，可以用计算机代数系统做这种计算。） ■

例 11 把 8 块相同的饼干分给 3 个不同的孩子，如果每个孩子至少接受 2 块饼干并且不超过 4 块饼干，那么有多少种不同的分配方式？

解 因为每个孩子至少接受 2 块饼干且不超过 4 块饼干，在关于序列 $\{C_n\}$ 的生成函数中对每个孩子存在一个等于

$$(x^2 + x^3 + x^4)$$

的因式，其中 c_n 是分配 n 块饼干的方式数。因为存在 3 个孩子，这个生成函数是

$$(x^2 + x^3 + x^4)^3$$

我们要求这个乘积中的 x^8 的系数。理由就是在展开式中 x^8 的项对应于选 3 项的方式数，其中每个因式选 1 项且指数加起来等于 8。此外，来自第一、第二和第三个因式的项的指数分别是第一、第二和第三个孩子接受的饼干数。通过计算说明这个系数等于 6。于是存在 6 种方式分配饼干使得每个孩子至少接受 2 块，但是不超过 4 块饼干。 ■

例 12 把价值 1 美元、2 美元和 5 美元的代币插入售货机为价值 r 美元的某种物品付款，使用生成函数确定在代币插入是有序的和无序的两种情况下付款的方式数。（例如为一种价值 3 美元的物品付款，当不考虑代币插入的次序时存在 2 种方式：插入 3 个 1 美元的代币或 1 个 1 美元和 1 个 2 美元的代币。当考虑代币插入的次序时有 3 种方式：插入 3 个 1 美元的代币，插入 1 个 1 美元代币然后 1 个 2 美元的代币，插入 1 个 2 美元代币然后 1 个 1 美元代币。）

解 在不考虑代币插入次序的情况下，我们所关心的就是为产生 r 美元的总数所使用的每种代币的数目。因为可以使用任意多个 1 美元的代币、任意多个 2 美元的代币和任意多个 5 美元的代币，答案就是在生成函数

$$(1 + x + x^2 + x^3 + \cdots)(1 + x^2 + x^4 + x^6 + \cdots)(1 + x^5 + x^{10} + x^{15} + \cdots)$$

中的 x^r 的系数。（这个乘积中的第一个因式表示所使用的 1 美元代币，第二个表示所使用的 2 美元代币，第三个表示所使用的 5 美元代币。）例如，用 1 美元、2 美元和 5 美元为一个价值 7 美元的物品付款的方式数由展开式中 x^7 的系数给出，结果等于 6。

当考虑代币插入的次序时，插入恰好 n 个代币产生 r 美元总数的方式数是在

$$(x + x^2 + x^5)^n$$

中的 x^r 的系数，因为这 n 个代币中的每一个可能是 1 美元代币、2 美元代币或 5 美元代币。又由于可以插入的代币不限数量，因此当考虑代币插入的次序时，使用 1 美元、2 美元或 5 美元代币产生 r 美元的方式数是在

$$1 + (x + x^2 + x^5) + (x + x^2 + x^5)^2 + \cdots = \frac{1}{1 - (x + x^2 + x^5)}$$

$$=\frac{1}{1-x-x^2-x^5}$$

中 x^r 的系数。这里我们把插入 0 个代币、1 个代币、2 个代币、3 个代币等方式数加起来，同时我们使用了恒等式 $1/(1-x)=1+x+x^2+\cdots$ 且用 $x+x^2+x^5$ 代替了 x。例如，用 1 美元、2 美元和 5 美元的代币为一个价值 7 美元的物品付款，当考虑使用代币的次序时，方式数是这个展开式中 x^7 的系数，等于 26。[为看到这个系数等于 26，要把 $(x+x^2+x^5)^k$ 的展开式中 x^7 的系数加起来，其中 $2\leqslant k\leqslant 7$。这项工作可以用大量的手工计算完成，也可以使用一个计算机代数系统来完成。] ■

例 13 说明了当求解带不同假设的问题时生成函数具有的多功能性。

例 13 假设已经建立了二项式定理，使用生成函数找出 n 元素集合的 k 组合数。

解 集合的 n 个元素中的每一个元素都对生成函数 $f(x)=\sum_{k=0}^{n}a_kx^k$ 贡献了项 $(1+x)$。因此 $f(x)$ 是关于 $\{a_k\}$ 的生成函数，其中 a_k 表示 n 元素集合的 k 组合数。于是，

$$f(x)=(1+x)^n$$

但是由二项式定理，我们有

$$f(x)=\sum_{k=0}^{n}\binom{n}{k}x^k$$

其中

$$\binom{n}{k}=\frac{n!}{k!(n-k)!}$$

于是，$C(n,k)$，n 元素集合的 k 组合数是

$$\frac{n!}{k!(n-k)!}$$ ■

注意 在 4.4 节我们使用了关于 n 元素集合的 r 组合数的公式证明了二项式定理。这些例子说明也可以用数学归纳法证明二项式定理，再用二项式定理推导关于 n 元素集合的 r 组合数的公式。

例 14 使用生成函数找出当元素允许重复时 n 元素集合的 r 组合数公式。

解 设 $G(x)$ 是关于序列 $\{a_r\}$ 的生成函数，其中 a_r 等于 n 元素集合的允许重复的 r 组合数，$G(x)=\sum_{r=0}^{\infty}a_rx^r$，即当我们构成允许重复的 r 组合时，对 n 元素集合的元素的选择不受限制，这 n 个元素中的每一个元素都对 $G(x)$ 的乘积展开式贡献了 $(1+x+x^2+x^3+\cdots)$ 这个因式。这是由于当一个 r 组合被构成时(要选总共 r 个元素)，每个元素都可以被选择 0 次、1 次、2 次、3 次，等等。因为集合中存在 n 个元素，且每一个都对 $G(x)$ 贡献了相同的因式，从而有

$$G(x)=(1+x+x^2+\cdots)^n$$

只要 $|x|<1$，我们就有 $1+x+x^2+\cdots=1/(1-x)$，所以

$$G(x)=1/(1-x)^n=(1-x)^{-n}$$

使用广义二项式定理(定理 2)，得到

$$(1-x)^{-n}=(1+(-x))^{-n}=\sum_{r=0}^{\infty}\binom{-n}{r}(-x)^r$$

当 r 是正整数时，n 元素集合的允许重复的 r 组合数就是这个和式中的 x^r 的系数。因此，使用例 8 我们求出 a_r 等于

$$\binom{-n}{r}(-1)^r=(-1)^r C(n+r-1,r)\cdot(-1)^r$$
$$=C(n+r-1,r)$$
■

注意例 14 的结果与我们在 4.5 节定理 2 所叙述的结果一样。

例 15 使用生成函数求出从 n 类不同的物体中选择 r 个物体并且每类物体至少选 1 个的方式数。

解 因为我们需要每类物体至少选 1 个，这 n 个类中的每类物体都对序列$\{a_r\}$的生成函数 $G(x)$贡献了因式$(1+x+x^2+x^3+\cdots)$其中 a_r 是从 n 类不同的物体中选择 r 个物体并且每类物体至少选 1 个的方式数。因此，

$$G(x)=(x+x^2+x^3+\cdots)^n=x^n(1+x+x^2+\cdots)^n=x^n/(1-x)^n$$

使用广义二项式定理和例 8，有

$$\begin{aligned}G(x)&=x^n/(1-x)^n\\&=x^n\cdot(1-x)^{-n}\\&=x^n\sum_{r=0}^{\infty}\binom{-n}{r}(-x)^r\\&=x^n\sum_{r=0}^{\infty}(-1)^r C(n+r-1,r)(-1)^r x^r\\&=\sum_{r=0}^{\infty}C(n+r-1,r)x^{n+r}\\&=\sum_{t=n}^{\infty}C(t-1,t-n)x^t\\&=\sum_{r=n}^{\infty}C(r-1,r-n)x^r\end{aligned}$$

在倒数第二个等式我们令 $t=n+r$ 以使得当 $r=0$ 时 $t=n$ 且 $n+r-1=t-1$，从而对求和进行了移位，然后在最后的等式中用 r 替换 t 作为和的下标而回到了初始的记号。因此，从 n 类不同的物体中选择 r 个物体，如果每类物体必须至少选 1 个时存在 $C(r-1,\ r-n)$种方式。 ■

6.4.4 使用生成函数求解递推关系

我们可以通过找相关的生成函数的显式公式来求解关于递推关系和初始条件的解。这可以用例 16 和例 17 来说明。

例 16 求解递推关系 $a_k=3a_{k-1}$，$k=1, 2, 3, \cdots$且初始条件 $a_0=2$。

Extra Examples **解** 设 $G(x)$是序列$\{a_k\}$的生成函数，即 $G(x)=\sum_{k=0}^{\infty}a_k x^k$。首先注意到

$$xG(x)=\sum_{k=0}^{\infty}a_k x^{k+1}=\sum_{k=1}^{\infty}a_{k-1}x^k$$

使用递推关系有

$$\begin{aligned}G(x)-3xG(x)&=\sum_{k=0}^{\infty}a_k x^k-3\sum_{k=1}^{\infty}a_{k-1}x^k\\&=a_0+\sum_{k=1}^{\infty}(a_k-3a_{k-1})x^k\\&=2\end{aligned}$$

因为 $a_0=2$ 且 $a_k=3a_{k-1}$，于是

$$G(x)-3xG(x)=(1-3x)G(x)=2$$

求解 $G(x)$ 得 $G(x)=2/(1-3x)$。使用表 6-1 的恒等式 $1/(1-ax)=\sum_{k=0}^{\infty} a^k x^k$，我们有

$$G(x)=2\sum_{k=0}^{\infty}3^k x^k=\sum_{k=0}^{\infty}2\cdot 3^k x^k$$

于是，$a_k=2\cdot 3^k$。■

例 17 设一个有效的编码字是一个包含偶数个 0 的十进制数字串。令 a_n 表示 n 位有效编码字的个数。在 6.1 节的例 7 中我们证明了序列 $\{a_n\}$ 满足递推关系

$$a_n=8a_{n-1}+10^{n-1}$$

且初始条件 $a_1=9$。使用生成函数找出关于 a_n 的显式公式。

解 为了简化关于生成函数的推导，我们通过置 $a_0=1$ 将序列扩充，当把这个值赋给 a_0 并且使用递推关系就得到 $a_1=8a_0+10^0=8+1=9$，这与初始条件一致。(由于存在一个长为 0 的编码字——空串，这也是有意义的。)

用 x^n 乘递推关系的两边得

$$a_n x^n=8a_{n-1}x^n+10^{n-1}x^n$$

设 $G(x)=\sum_{n=0}^{\infty}a_n x^n$ 是序列 $a_0, a_1, a_2, \cdots$ 的生成函数，从 $n=1$ 开始对上面的等式两边求和，得到

$$\begin{aligned}G(x)-1&=\sum_{n=1}^{\infty}a_n x^n=\sum_{n=1}^{\infty}(8a_{n-1}x^n+10^{n-1}x^n)\\&=8\sum_{n=1}^{\infty}a_{n-1}x^n+\sum_{n=1}^{\infty}10^{n-1}x^n\\&=8x\sum_{n=1}^{\infty}a_{n-1}x^{n-1}+x\sum_{n=1}^{\infty}10^{n-1}x^{n-1}\\&=8x\sum_{n=0}^{\infty}a_n x^n+x\sum_{n=0}^{\infty}10^n x^n\\&=8xG(x)+x/(1-10x)\end{aligned}$$

其中我们已经使用了例 5 对第二个和进行求值。因此有

$$G(x)-1=8xG(x)+x/(1-10x)$$

求解 $G(x)$ 得

$$G(x)=\frac{1-9x}{(1-8x)(1-10x)}$$

把等式的右边展开成部分分式(正如在微积分中研究有理函数的积分时所做的)得到

$$G(x)=\frac{1}{2}\left(\frac{1}{1-8x}+\frac{1}{1-10x}\right)$$

两次使用例 5(一次设 $a=8$，一次设 $a=10$)得

$$\begin{aligned}G(x)&=\frac{1}{2}\left(\sum_{n=0}^{\infty}8^n x^n+\sum_{n=0}^{\infty}10^n x^n\right)\\&=\sum_{n=0}^{\infty}\frac{1}{2}(8^n+10^n)x^n\end{aligned}$$

于是，证明了

$$a_n=\frac{1}{2}(8^n+10^n)$$

■

6.4.5 使用生成函数证明恒等式

在第4章我们已经看到怎样使用组合证明方法来建立组合恒等式。这里将显示这种恒等式，还有关于广义二项式系数的恒等式，都可以使用生成函数来证明。有时候生成函数的方法比其他方法更简单，特别是用生成函数的封闭形式比使用序列本身更能简化证明过程。我们用例18说明怎样用生成函数证明恒等式。

例18 使用生成函数证明

$$\sum_{k=0}^{n} C(n,k)^2 = C(2n,n)$$

其中 n 是正整数。

解 首先注意到根据二项式定理 $C(2n, n)$ 是 $(1+x)^{2n}$ 中 x^n 的系数。而我们也有

$$\begin{aligned}(1+x)^{2n} &= [(1+x)^n]^2 \\ &= [C(n,0)+C(n,1)x+C(n,2)x^2+\cdots+C(n,n)x^n]^2\end{aligned}$$

在这个展开式中 x^n 的系数是 $C(n, 0)C(n, n)+C(n, 1)C(n, n-1)+C(n, 2)C(n, n-2)+\cdots+C(n, n)C(n, 0)$。因为 $C(n, n-k)=C(n, k)$，它等于 $\sum_{k=0}^{n} C(n,k)^2$。由于 $C(2n,n)$ 和 $\sum_{k=0}^{n} C(n,k)^2$ 都表示 $(1+x)^{2n}$ 中 x^n 的系数，它们一定是相等的。 ■

本节末的练习42和43要求用生成函数来证明帕斯卡恒等式和范德蒙德恒等式。

练习

1. 求关于有穷序列2，2，2，2，2，2的生成函数。
2. 求关于有穷序列1，4，16，64，256的生成函数。
3. 求关于下面每个序列生成函数的封闭形式。(用最明显的选择设定每个序列的通项形式。)
   a) 0，2，2，2，2，2，2，0，0，0，0，0，…
   b) 0，0，0，1，1，1，1，1，1，…
   c) 0，1，0，0，1，0，0，1，0，0，1，…
   d) 2，4，8，16，32，64，128，256，…
   e) $\binom{7}{0}$，$\binom{7}{1}$，$\binom{7}{2}$，…，$\binom{7}{7}$，0，0，0，0，0，…
   f) 2，−2，2，−2，2，−2，2，−2，…
   g) 1，1，0，1，1，1，1，1，1，1，…
   h) 0，0，0，1，2，3，4，…
4. 求关于下面每个序列生成函数的封闭形式。(用最明显的选择设定每个序列的通项形式。)
   a) −1，−1，−1，−1，−1，−1，−1，0，0，0，0，0，0，…
   b) 1，3，9，27，81，243，729，…
   c) 0，0，3，−3，3，−3，3，−3，…
   d) 1，2，1，1，1，1，1，1，1，…
   e) $\binom{7}{0}$，$2\binom{7}{1}$，$2^2\binom{7}{2}$，…，$2^7\binom{7}{7}$，0，0，0，0，…
   f) −3，3，−3，3，−3，3，…
   g) 0，1，−2，4，−8，16，−32，64，…
   h) 1，0，1，0，1，0，1，0，…
5. 求关于序列 $\{a_n\}$ 的生成函数的封闭形式，其中
   a) $a_n=5$，对所有的 $n=0$，1，2…
   b) $a_n=3^n$，对所有的 $n=0$，1，2…
   c) $a_n=2$，对 $n=3$，4，5，…且 $a_0=a_1=a_2=0$
   d) $a_n=2n+3$ 对所有的 $n=0$，1，2，…

e) $a_n=\binom{8}{n}$对所有的 $n=0$，1，2…　　f) $a_n=\binom{n+4}{n}$对所有的 $n=0$，1，2…

6. 求关于序列$\{a_n\}$的生成函数的封闭形式，其中

a) $a_n=-1$，对所有的 $n=0$，1，2…　　b) $a_n=2^n$，对 $n=1$，2，3，4，…且 $a_0=0$

c) $a_n=n-1$，对 $n=0$，1，2，…　　d) $a_n=1/(n+1)!$，对 $n=0$，1，2，…

e) $a_n=\binom{n}{2}$，对 $n=0$，1，2…　　f) $a_n=\binom{10}{n+1}$，对 $n=0$，1，2…

7. 对于下面每一个生成函数给出关于它所确定序列的封闭公式。

a) $(3x-4)^3$　　b) $(x^3+1)^3$

c) $1/(1-5x)$　　d) $x^3/(1+3x)$

e) $x^2+3x+7+(1/(1-x^2))$　　f) $(x^4/(1-x^4))-x^3-x^2-x-1$

g) $x^2/(1-x)^2$　　h) $2e^{2x}$

8. 对于下面每一个生成函数给出关于它所确定序列的封闭公式。

a) $(x^2+1)^3$　　b) $(3x-1)^3$　　c) $1/(1-2x^2)$　　d) $x^2/(1-x)^3$

e) $x-1+(1/(1-3x))$　　f) $(1+x^3)/(1+x)^3$　　*g) $x/(1+x+x^2)$　　h) $e^{3x^2}-1$

9. 求出下面每个函数的幂级数中 x^{10} 的系数。

a) $(1+x^5+x^{10}+x^{15}+\cdots)^3$

b) $(x^3+x^4+x^5+x^6+x^7+\cdots)^3$

c) $(x^4+x^5+x^6)(x^3+x^4+x^5+x^6+x^7)(1+x+x^2+x^3+x^4+\cdots)$

d) $(x^2+x^4+x^6+x^8+\cdots)(x^3+x^6+x^9+\cdots)(x^4+x^8+x^{12}+\cdots)$

e) $(1+x^2+x^4+x^6+x^8+\cdots)(1+x^4+x^8+x^{12}+\cdots)(1+x^6+x^{12}+x^{18}+\cdots)$

10. 求出下面每个函数的幂级数中 x^9 的系数。

a) $(1+x^3+x^6+x^9+\cdots)^3$

b) $(x^2+x^3+x^4+x^5+x^6+\cdots)^3$

c) $(x^3+x^5+x^6)(x^3+x^4)(x+x^2+x^3+x^4+\cdots)$

d) $(x+x^4+x^7+x^{10}+\cdots)(x^2+x^4+x^6+x^8+\cdots)$

e) $(1+x+x^2)^3$

11. 求出下面每个函数的幂级数中 x^{10} 的系数。

a) $1/(1-2x)$　　b) $1/(1+x)^2$　　c) $1/(1-x)^3$　　d) $1/(1+2x)^4$　　e) $x^4/(1-3x)^3$

12. 求出下面每个函数的幂级数中 x^{12} 的系数。

a) $1/(1+3x)$　　b) $1/(1-2x)^2$　　c) $1/(1+x)^8$　　d) $1/(1-4x)^3$　　e) $x^3/(1+4x)^2$

13. 把 10 个相同的球分给 4 个孩子，如果每个孩子至少得到 2 个球，使用生成函数确定不同的分法数。

14. 把 12 张相同的剧情图片分给 5 个孩子使得每个孩子至多得到 3 张，使用生成函数确定不同的分法数。

15. 把 15 个相同的动物玩具分给 6 个孩子使得每个孩子至少得到 1 个但不超过 3 个，使用生成函数确定不同的分法数。

16. 从 3 类百吉饼——鸡蛋的、椒盐的和普通的选 12 个，如果每类至少选 2 个但椒盐的不超过 3 个，使用生成函数确定选法数。

17. 把 25 个相同的多纳圈分给 4 个警官使得每个警官至少得到 3 个但不超过 7 个，有多少种方式？

18. 从包含 100 个红球、100 个蓝球和 100 个绿球的罐子中选 14 个球，使得蓝球不少于 3 个且不多于 10 个。假定不考虑选球的次序，使用生成函数求出选法数。

19. 求序列$\{c_k\}$的生成函数，其中 c_k 是使用 1 美元、2 美元、5 美元和 10 美元纸币换 k 美元的方法数。

20. 求序列$\{c_k\}$的生成函数，其中 c_k 是使用 10 比索、20 比索、50 比索和 100 比索换 k 比索的方法数。

21. 对$(1+x+x^2+x^3+\cdots)^3$ 展开式中 x^4 的系数给出组合解释。使用这个解释求出这个数。

22. 对$(1+x+x^2+x^3+\cdots)^n$展开式中x^6的系数给出组合解释。使用这个解释求出这个数。

23. a)什么是关于$\{a_k\}$的生成函数？这里的a_k是$x_1+x_2+x_3=k$的解的个数，其中x_1，x_2和x_3是满足$x_1\geqslant 2$，$0\leqslant x_2\leqslant 3$，$2\leqslant x_3\leqslant 5$的整数。

    b)使用(a)的答案求a_6。

24. a)什么是关于$\{a_k\}$的生成函数？这里的a_k是$x_1+x_2+x_3+x_4=k$的解的个数，其中x_1，x_2，x_3和x_4是满足$x_1\geqslant 3$，$1\leqslant x_2\leqslant 5$，$0\leqslant x_3\leqslant 4$，$x_4\geqslant 1$的整数。

    b)使用(a)的答案求a_7。

25. 解释怎样使用生成函数找到用 3 分、4 分和 20 分的邮票在信封上贴满 r 分邮费的方式数。

    a)假设不考虑贴邮票的次序。

    b)假设邮票贴成一行并且考虑贴的次序。

    c)当不考虑贴邮票的次序时，使用(a)的答案确定用 3 分、4 分和 20 分的邮票在信封上贴满 46 分邮费的方式数。(建议使用计算机代数程序。)

    d)当考虑贴邮票的次序时，使用(b)的答案确定用 3 分、4 分和 20 分的邮票在信封上贴满一行 46 分邮费的方式数。(建议使用计算机代数程序。)

26. a)重复掷一个骰子，考虑掷的次序并且使得掷出的点数之和为 n，证明关于这种方式数的生成函数是$1/(1-x-x^2-x^3-x^4-x^5-x^6)$。

    b)使用(a)求出重复掷一个骰子，考虑掷的次序并且使得掷出的总点数为 8 的方式数。(建议使用计算机代数程序。)

27. 使用生成函数(如果需要的话，也使用计算机代数程序)求出换 1 美元的方式数。

    a)用 10 美分和 25 美分。

    b)用 5 美分、10 美分和 25 美分。

    c)用 1 美分、10 美分和 25 美分。

    d)用 1 美分、5 美分、10 美分和 25 美分。

28. 使用生成函数(如果需要的话，也使用计算机代数程序)求出用 1 美分、5 美分、10 美分和 25 美分换 1 美元的方式数，使得

    a)1 美分不超过 10 个。

    b)1 美分不超过 10 个且 5 美分不超过 10 个。

    *c)硬币不超过 10 个。

29. 使用生成函数求出换 100 美元的方式数。

    a)用 10 美元、20 美元和 50 美元纸币。

    b)用 5 美元、10 美元、20 美元和 50 美元纸币。

    c)用 5 美元、10 美元、20 美元和 50 美元纸币，并且每种纸币至少使用 1 张。

    d)用 5 美元、10 美元和 20 美元纸币，并且每种纸币至少使用 1 张但不超过 4 张。

30. 如果$G(x)$是关于序列$\{a_k\}$的生成函数，那么关于下述每个序列的生成函数是什么？

    a)$2a_0, 2a_1, 2a_2, 2a_3, \cdots$

    b)$0, a_0, a_1, a_2, a_3, \cdots$

    c)$0, 0, 0, 0, a_2, a_3, \cdots$

    d)$a_2, a_3, a_4, \cdots$

    e)$a_1, 2a_2, 3a_3, 4a_4, \cdots$[提示：这里需要微积分。]

    f)$a_0^2, 2a_0a_1, a_1^2+2a_0a_2, 2a_0a_3+2a_1a_2, 2a_0a_4+2a_1a_3+a_2^2, \cdots$

31. 如果$G(x)$是关于序列$\{a_k\}$的生成函数，那么关于下述每个序列的生成函数是什么？

    a)$0, 0, 0, a_3, a_4, a_5, \cdots$

    b)$a_0, 0, a_1, 0, a_2, 0, \cdots$

    c)$0, 0, 0, 0, a_0, a_1, a_2, \cdots$

    d)$a_0, 2a_1, 4a_2, 8a_3, 16a_4, \cdots$

    e)$a_0, a_1/2, a_2/3, a_3/4, \cdots$[提示：这里需要微积分。]

    f)$a_0, a_0+a_1, a_0+a_1+a_2, a_0+a_1+a_2+a_3, \cdots$

32. 使用生成函数求解递推关系$a_k=7a_{k-1}$，初始条件$a_0=5$。

33. 使用生成函数求解递推关系$a_k=3a_{k-1}+2$，初始条件$a_0=1$。

34. 使用生成函数求解递推关系$a_k=3a_{k-1}+4^{k-1}$，初始条件$a_0=1$。

35. 使用生成函数求解递推关系 $a_k=5a_{k-1}-6a_{k-2}$，初始条件 $a_0=6$ 和 $a_1=30$。

36. 使用生成函数求解递推关系 $a_k=a_{k-1}+2a_{k-2}+2^k$，初始条件 $a_0=4$ 和 $a_1=12$。

37. 使用生成函数求解递推关系 $a_k=4a_{k-1}-4a_{k-2}+k^2$，初始条件 $a_0=2$ 和 $a_1=5$。

38. 使用生成函数求解递推关系 $a_k=2a_{k-1}+3a_{k-2}+4^k+6$，初始条件 $a_0=20$ 和 $a_1=60$。

39. 使用生成函数找出关于斐波那契数的显式公式。

*40. a)证明如果 n 是正整数，那么

$$\binom{-1/2}{n}=\frac{\binom{2n}{n}}{(-4)^n}$$

b)使用广义二项式定理和(a)证明对于一切非负整数 n，在$(1-4x)^{-1/2}$的展开式中 x^n 的系数是$\binom{2n}{n}$。

*41. (需要微积分)设$\{C_n\}$是卡特朗数的序列，即具有初值 $C_0=C_1=1$ 的递推关系 $C_n=\sum_{k=0}^{n-1}C_kC_{n-1-k}$ 的解(见 6.1 节例 8)。

a)证明如果 $G(x)$是关于卡特朗数的序列的生成函数，那么 $xG(x)^2-G(x)+1=0$。(使用初始条件)断定 $G(x)=(1-\sqrt{1-4x})/(2x)$。

b)使用练习 40 断定

$$G(x)=\sum_{n=0}^{\infty}\frac{1}{n+1}\binom{2n}{n}x^n$$

从而

$$C_n=\frac{1}{n+1}\binom{2n}{n}$$

42. 当 n 和 r 是 $r<n$ 的正整数时，使用生成函数证明帕斯卡恒等式：$C(n,r)=C(n-1,r)+C(n-1,r-1)$。[提示：使用恒等式$(1+x)^n=(1+x)^{n-1}+x(1+x)^{n-1}$。]

43. 使用生成函数证明范德蒙恒等式：$C(m+n,r)=\sum_{k=0}^{r}C(m,r-k)C(n,k)$，其中 m,n 和 r 是非负整数，且 r 不超过 m 或 n 。[提示：看在$(1+x)^{m+n}=(1+x)^m(1+x)^n$ 两边的 x^r 的系数。]

44. 这个练习说明了怎样使用生成函数推导前 n 个平方数之和的公式。

a)证明$(x^2+x)/(1-x)^4$ 是关于序列$\{a_n\}$的生成函数，其中 $a_n=1^2+2^2+\cdots+n^2$。

b)使用(a)找出关于和 $1^2+2^2+\cdots+n^2$ 的显式公式。

关于序列$\{a_n\}$的指数生成函数是级数

$$\sum_{n=0}^{\infty}\frac{a_n}{n!}x^n$$

例如，关于序列 1，1，1，…的指数生成函数是 $\sum_{n=0}^{\infty}x^n/n!=e^x$。(你将发现这个级数在下面的练习中很有用。)注意 e^x 是关于序列 1,1,1/2!,1/3!,1/4!,… 的普通生成函数。

45. 求一个关于序列$\{a_n\}$的指数生成函数的封闭形式，其中

a)$a_n=2$　　b)$a_n=(-1)^n$　　c)$a_n=3^n$

d)$a_n=n+1$　　e)$a_n=1/(n+1)$

46. 求一个关于序列$\{a_n\}$的指数生成函数的封闭形式，其中

a)$a_n=(-2)^n$　　b)$a_n=-1$　　c)$a_n=n$

d)$a_n=n(n-1)$　　e)$a_n=1/((n+1)(n+2))$

47. 求以下述函数为指数生成函数的序列。

a) $f(x)=e^{-x}$　b) $f(x)=3x^{2x}$　c) $f(x)=e^{3x}-3e^{2x}$

d) $f(x)=(1-x)+e^{-2x}$　e) $f(x)=e^{-2x}-(1/(1-x))$　f) $f(x)=e^{-3x}-(1+x)+(1/(1-2x))$

g) $f(x)=e^{x^2}$

48. 求以下述函数为指数生成函数的序列。

a) $f(x)=e^{3x}$　b) $f(x)=2e^{-3x+1}$　c) $f^{(x)}=e^{4x}+e^{-4x}$

d) $f(x)=(1+2x)+e^{3x}$　e) $f(x)=e^{x}-(1/1+x))$　f) $f(x)=xe^{x}$

g) $f(x)=e^{x^3}$

49. 一个编码系统用八进制(基为8)数字串对信息编码。一个编码字是有效的，当且仅当它包含偶数个7。

a)求一个关于n位长有效编码字个数的线性非齐次递推关系。初始条件是什么?

b)使用6.2节的定理6解这个递推关系。

c)用生成函数解这个递推关系。

*50. 一个编码系统用四进制数字串(即数字来自集合{0，1，2，3})对信息编码。一个编码字是有效的当且仅当它包含偶数个0和偶数个1。设a_n等于长为n的有效编码字个数。此外令b_n为具有偶数个0和奇数个1的n位四进制数字串个数，c_n为具有奇数个0和偶数个1的n位四进制数字串个数，d_n为具有奇数个0和奇数个1的n位四进制数字串个数。

a)证明$d_n=4^n-a_n-b_n-c_n$。使用这个式子证明$a_{n+1}=2a_n+b_n+c_n$，$b_{n+1}=b_n-c_n+4^n$和$c_{n+1}=c_n-b_n+4^n$。

b) a_1，b_1，c_1和d_1是什么?

c)使用(a)和(b)求出a_3，b_3，c_3和d_3。

d)使用(a)的递推关系和(b)的初始条件分别建立与序列$\{a_n\}$，$\{b_n\}$和$\{c_n\}$的生成函数$A(x)$，$B(x)$和$C(x)$相关的三个方程。

e)求解(d)的方程得到关于$A(x)$，$B(x)$和$C(x)$的显式公式，并且利用这些公式得到关于a_n，b_n，c_n和d_n的显式公式。

在研究整数n的不同类型的剖分个数时生成函数是很有用的。一个正整数的剖分是把这个整数写成正整数之和，和中的整数允许重复并且不考虑次序。例如，5的剖分(不加限制)是1+1+1+1+1，1+1+1+2，1+1+3，1+2+2，1+4，2+3和5。练习51～56说明了这种应用。

51. 证明在$1/((1-x)(1-x^2)(1-x^3)\cdots)$的形式幂级数展开式中$x^n$的系数$p(n)$等于$n$的剖分数。

52. 证明在$1/((1-x)(1-x^3)(1-x^5)\cdots)$的形式幂级数展开式中$x^n$的系数$p_o(n)$等于$n$剖分成奇整数(即把$n$写成正奇数之和)的方式数，其中不管这些奇数的次序并且允许重复。

53. 证明在$(1+x)(1+x^2)(1+x^3)\cdots$的形式幂级数展开式中$x^n$的系数$p_d(n)$等于$n$剖分成不相等的整数(即把$n$写成正整数之和)的方式数，其中不管这些整数的次序但不允许重复。

54. 对于$1\leqslant n\leqslant 8$，通过对每个整数写出每一个不同类型的剖分求$p_o(n)$和$p_d(n)$，其中$p_o(n)$是n剖分成允许重复的奇整数的方式数，$p_d(n)$是n剖分成不相等的整数的方式数。

55. 证明如果n是正整数，那么n剖分成不相等的整数的方式数等于n剖分成允许重复的奇整数的方式数；即$p_o(n)=p_d(n)$。[提示：证明关于$p_o(n)$和$p_d(n)$的生成函数相等。]

**56. (需要微积分)使用关于$p(n)$的生成函数证明对某个常数C，$p(n)\leqslant e^{C\sqrt{n}}$。[Hardy和Ramanujan证明了$p(n)\approx e^{\pi\sqrt{2/3}\sqrt{n}}/(4\sqrt{3}n)$，这意味着当$n$达到无限时$p(n)$与右边的比达到1。]

假定X是样本空间S上的随机变量，使得$X(s)$对于所有的$s\in S$是非负整数。关于X的概率生成函数是

$$G_X(x)=\sum_{k=0}^{\infty}p(X(s)=k)x^k$$

57. (需要微积分)证明如果G_X是随机变量X的概率生成函数，使得$X(s)$对于所有的$s\in S$是非负整数，

那么

a) $G_X(1)=1$ b) $E(x)=G'_X(1)$ c) $V(X)=G''_X(1)+G'_X(1)-G'_X(1)^2$

58. 做独立的伯努利实验，每次实验成功的概率为 p。设 X 是随机变量，且如果第 n 次实验出现首次成功，X 的值就是 n。

a)求关于概率生成函数 G_X 的封闭公式。

b)使用练习 57 和(a)中得到的关于概率生成函数的封闭公式求 X 的期望值和方差。

59. 设 m 是正整数，当做独立的伯努利实验时每次实验成功的概率为 p。设 X_m 是随机变量，且如果第 $n+m$ 次实验出现第 m 次成功则 X 的值就是 n。

a)使用第 5 章的补充练习 30 证明概率生成函数 G_{X_m} 由 $G_{X_m}(x)=p^m/(1-qx)^m$ 给出，其中 $q=1-p$。

b)使用练习 57 和(a)中得到的关于概率生成函数的封闭公式求 X_m 的期望值和方差。

60. 证明如果 X 和 Y 是样本空间 S 上的独立的随机变量，使得 $X(s)$ 和 $Y(s)$ 对于所有 $s\in S$ 为非负整数，那么 $G_{X+Y}(x)=G_X(x)G_Y(x)$。

6.5 容斥

6.5.1 引言

一个离散数学班包含 30 个女生和 50 个二年级学生。在这个班里有多少个女生或二年级学生？如果没有更多的信息，这个问题是没法求解的。把女生数和二年级学生数加起来不一定能得出正确的结果，因为二年级的女生可能被计数了两次。这个事实说明，在班里的女生或二年级学生数是班里的女生数与二年级学生数之和减去二年级的女生数。在 4.1 节曾经介绍过求解这种计数问题的技术。这里我们将把在那一节引入的思想加以推广，以求解更为广泛的计数问题。

6.5.2 容斥原理

两个有穷集的并集中存在多少个元素？在 1.7 节中证明了两个集合 A 和 B 的并集中的元素数是这些集合的元素数之和减去其交集中的元素数，即

$$|A\cup B|=|A|+|B|-|A\cap B|$$

正如我们在 4.1 节证明的，这个关于两个集合并集中元素数的公式在计数问题中是很有用的。例 1～例 3 进一步说明了这个公式的用处。

例 1 一个离散数学班包含 25 个计算机科学专业的学生、13 个数学专业的学生和 8 个同时主修数学和计算机科学两个专业的学生。如果每个学生主修数学专业、计算机科学专业或者同时主修这两个专业，那么班里有多少个学生？

解 设 A 是这个班里计算机科学专业的学生的集合，B 是这个班里数学专业的学生的集合，那么 $A\cap B$ 是班里主修数学和计算机科学两个专业的学生的集合。因为这个班的每个学生或者主修计算机科学，或者主修数学（或者同时主修两个专业），从而得到这个班里的学生数是 $|A\cup B|$。于是

$$\begin{aligned}|A\cup B|&=|A|+|B|-|A\cap B|\\&=25+13-8\\&=30\end{aligned}$$

因此，这个班有 30 个学生。这个计算在图 6-7 中说明。 ■

例 2 有多少个不超过 1000 的正整数可以被 7 或 11 整除？

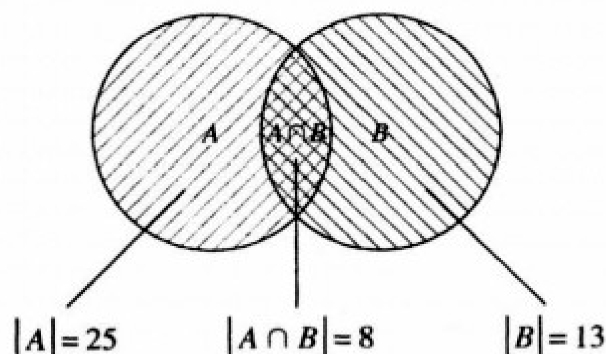


图 6-7 离散数学班的学生的集合

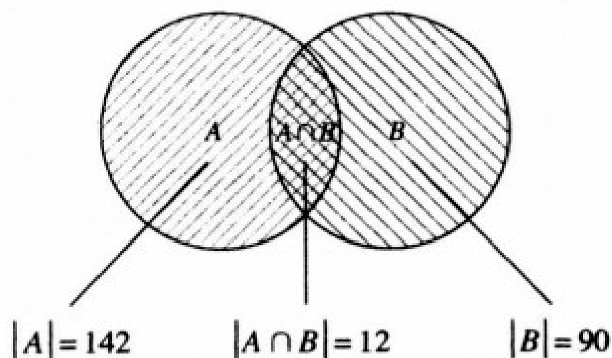


图 6-8 不超过 1000 的可被 7 或 11 整除的正整数的集合

解 设 A 是不超过 1000 且可被 7 整除的正整数的集合，B 是不超过 1000 且可被 11 整除的正整数的集合，那么 $A\cup B$ 是不超过 1000 且可被 7 或 11 整除的正整数的集合，$A\cap B$ 是不超过 1000 且可被 7 和 11 同时整除的正整数的集合。由 2.4 节的例 2，我们知道在不超过 1000 的正整数中有 $\lfloor 1000/7\rfloor$ 个整数可被 7 整除，并且有 $\lfloor 1000/11\rfloor$ 个整数可被 11 整除。由于 7 和 11 是互素的，被 7 和 11 同时整除的整数就是被 7 · 11 整除的整数。因此，有 $\lfloor 1000/(7\cdot 11)\rfloor$ 个不超过 1000 的正整数可被 7 和 11 同时整除。于是存在

$$
\begin{aligned}
|A\cup B| &= |A|+|B|-|A\cap B| \\
&= \left\lfloor\frac{1000}{7}\right\rfloor+\left\lfloor\frac{1000}{11}\right\rfloor-\left\lfloor\frac{1000}{7\cdot 11}\right\rfloor \\
&= 142+90-12 \\
&= 220
\end{aligned}
$$

个正整数不超过 1000 且可被 7 或 11 整除。这个计算在图 6-8 中说明。 ■

例 3 说明怎样求有穷全集中在两个集合的并集之外的元素数。

例 3 假设你们学校有 1807 个新生。其中 453 人选了一门计算机科学课，567 人选了一门数学课，299 人同时选了计算机科学课和数学课。有多少学生既没有选计算机科学课也没有选数学课？

解 为找出既没有选数学课也没有选计算机科学课的新生数，就要从新生总数中减去选了其中一门课的学生数。设 A 是选了一门计算机课的所有新生的集合，且 B 是选了一门数学课的所有新生的集合。于是 $|A|=453$，$|B|=567$，且 $|A\cap B|=299$。选了一门计算机科学或数学课的学生数是

$$
\begin{aligned}
|A\cup B| &= |A|+|B|-|A\cap B| \\
&= 453+567-299=721
\end{aligned}
$$

因此，存在 $1807-721=1086$ 个新生既没选计算机科学课也没选数学课。这个计算在图 6-9 中加以说明。 ■

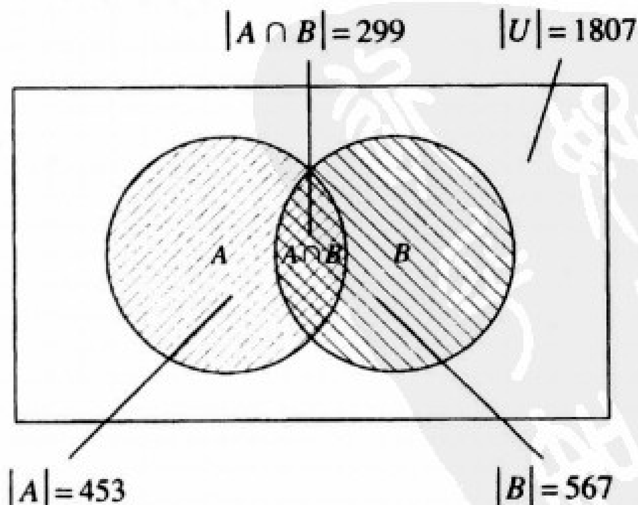
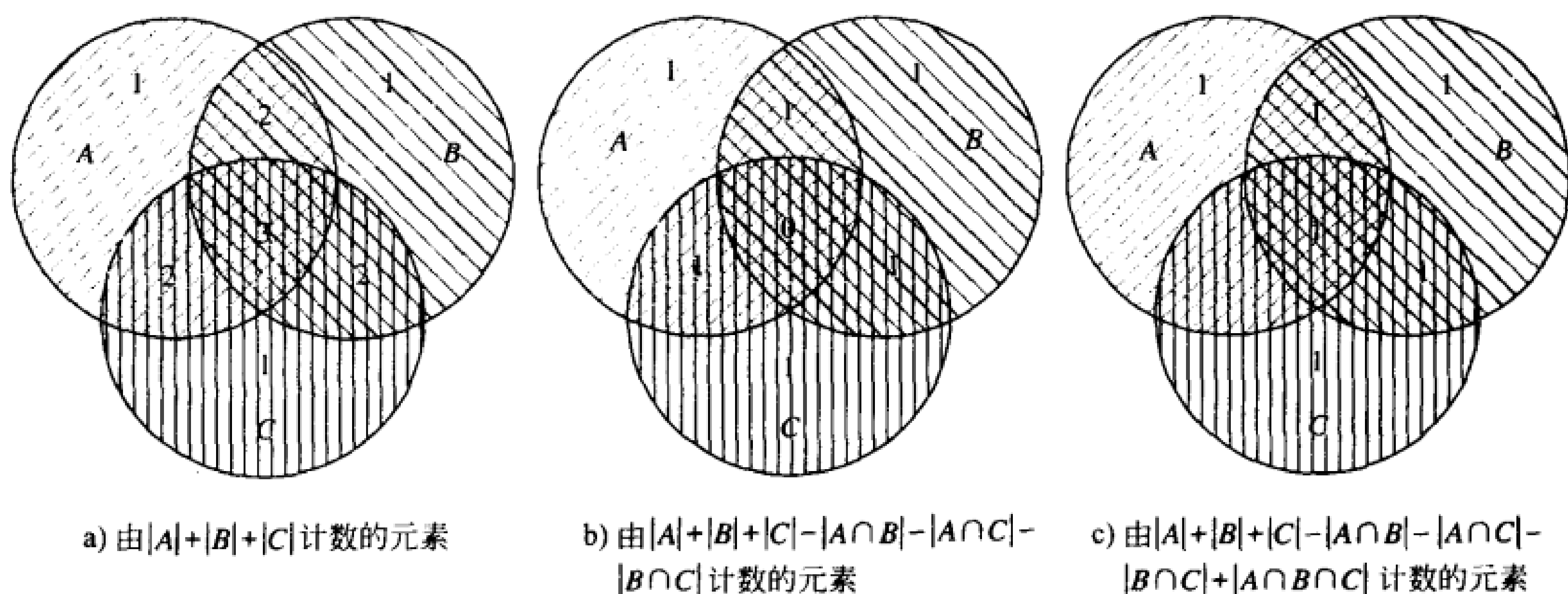


图 6-9 既没选计算机科学课也没选数学课的新生的集合

在本节的后面将说明怎样求有限个集合的并集中的元素数。这个结果叫做**容斥原理**。设 n 是任意正整数，在考虑 n 个集合的并集之前，先推导与 3 个集合 A，B 和 C 的并集中的元素数有关的公式。为推导这个公式，首先注意到以下事实：$|A|+|B|+$

$|C|$ 对 3 个集合中那些恰好在其中 1 个集合的元素只计数了 1 次，恰好在其中 2 个集合的元素计数了 2 次，恰好在其中 3 个集合的元素计数了 3 次。这个结果在图 6-10a 中说明。



a) 由 $|A|+|B|+|C|$ 计数的元素　b) 由 $|A|+|B|+|C|-|A\cap B|-|A\cap C|-|B\cap C|$ 计数的元素　c) 由 $|A|+|B|+|C|-|A\cap B|-|A\cap C|-|B\cap C|+|A\cap B\cap C|$ 计数的元素

图 6-10　求关于 3 个集合的并集中元素数的公式

为了去掉在多个集合中元素的重复计数，减去这 3 个集合中的每 2 个集合的交集中的元素数，得到

$$|A|+|B|+|C|-|A\cap B|-|A\cap C|-|B\cap C|$$

这个表达式对恰好出现在其中 1 个集合的元素仍旧计数 1 次。对恰好出现在其中 2 个集合的元素也计数 1 次，因为 2 个集合的交集有 3 个，而这种元素只出现在其中之一。但是，那些出现在 3 个集合的元素将被这个表达式计数 0 次，因为它们将会出现在所有的两两相交的 3 个交集中。这个结果显示在图 6-10b 中。

为了纠正这个漏计，还要加上 3 个集合交集中的元素数。这个最后的表达式对每个元素计数了 1 次，不管它是在 1 个、2 个还是在 3 个集合里。于是

$$|A\cup B\cup C|=|A|+|B|+|C|-|A\cap B|-|A\cap C|-|B\cap C|+|A\cap B\cap C|$$

这个公式显示在图 6-10c 中。

例 4 说明了怎样使用这个公式。

例 4　1232 个学生选了西班牙语课，879 个学生选了法语课，114 个学生选了俄语课。103 个学生选了西班牙语和法语课，23 个学生选了西班牙语和俄语课，14 个学生选了法语和俄语课。如果 2092 个学生至少在西班牙语、法语和俄语课中选 1 门，有多少个学生选了所有这 3 门语言课？

解　设 S 是选西班牙语课的学生集合，F 是选法语课的学生集合，R 是选俄语课的学生集合。那么

$$|S|=1232\quad |F|=879\quad |R|=114$$
$$|S\cap F|=103\quad |S\cap R|=23\quad |F\cap R|=14$$

且

$$|S\cup F\cup R|=2092$$

把这些等式代入下面的等式

$$|S\cup F\cup R|=|S|+|F|+|R|-|S\cap F|-|S\cap R|-|F\cap R|+|S\cap F\cap R|$$

得

$$2092=1232+879+114-103-23-14+|S\cap F\cap R|$$

求解上式得出 $|S\cap F\cap R|=7$。因此有 7 个学生同时选了西班牙语、法语和俄语课。这个结果在图 6-11 中说明。 ■

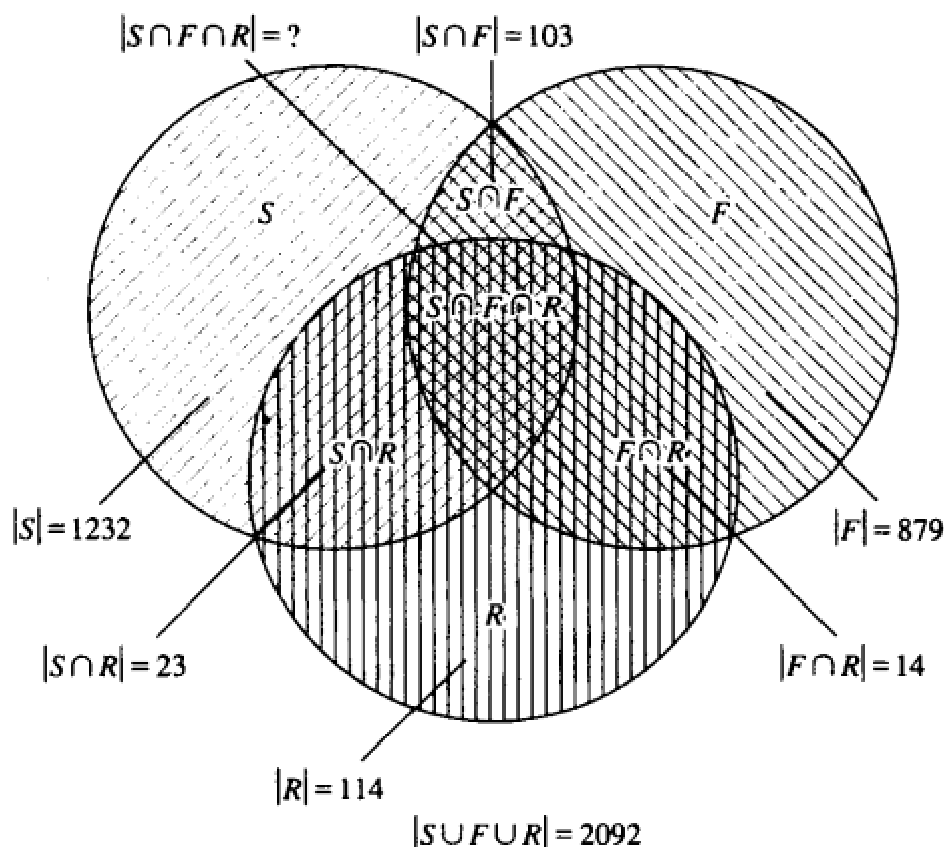


图 6-11 选了西班牙语、法语和俄语课程的学生集合

我们现在将叙述和证明容斥原理，它将告诉我们在有限个有穷集的并集中有多少个元素。

定理 1 容斥原理 设 $A_1, A_2, \cdots, A_n$ 是有穷集。那么

$$|A_1\cup A_2\cup\cdots\cup A_n|=\sum_{1\leqslant i\leqslant n}|A_i|-\sum_{1\leqslant i<j\leqslant n}|A_i\cap A_j|$$
$$+\sum_{1\leqslant i<j<k\leqslant n}|A_i\cap A_j\cap A_k|-\cdots+(-1)^{n+1}|A_1\cap A_2\cap\cdots\cap A_n|$$

证 我们将通过证明并集中的每个元素在等式右边恰好被计数 1 次来证明这个公式。假设 a 恰好是 $A_1, A_2, \cdots, A_n$ 中 r 个集合的成员，其中 $1\leqslant r\leqslant n$。这个元素被 $\sum|A_i|$ 计数了 $C(r,1)$ 次，被 $\sum|A_i\cap A_j|$ 计数了 $C(r,2)$ 次。一般说来，它被涉及 m 个 A_i 集合的求和计数了 $C(r,m)$ 次。于是，这个元素恰好被等式右边的表达式计数了

$$C(r,1)-C(r,2)+C(r,3)-\cdots+(-1)^{r+1}C(r,r)$$

次。我们的目标是求出这个值。由 4.4 节的推论 2，我们有

$$C(r,0)-C(r,1)+C(r,2)-\cdots+(-1)^rC(r,r)=0$$

于是

$$1=C(r,0)=C(r,1)-C(r,2)+\cdots+(-1)^{r+1}C(r,r)$$

因此，并集中的每个元素在等式右边的表达式中恰好被计数 1 次。这就证明了容斥原理。 □

对于每个正整数 n，容斥原理对于 n 个集合并集的元素数给出了一个公式。对于 n 个集合的集合族的每一个非空子集的交，在这个公式中都存在一项计数了它的元素。因此在这个公式中有 2^n-1 项。

例 5 对于 4 个集合的并集中的元素数给出一个公式。

解 容斥原理显示

$$|A_1\cup A_2\cup A_3\cup A_4|=|A_1|+|A_2|+|A_3|+|A_4|$$
$$-|A_1\cap A_2|-|A_1\cap A_3|-|A_1\cap A_4|-|A_2\cap A_3|-|A_2\cap A_4|-|A_3\cap A_4|$$

$$+|A_1\cap A_2\cap A_3|+|A_1\cap A_2\cap A_4|+|A_1\cap A_3\cap A_4|+|A_2\cap A_3\cap A_4|$$
$$-|A_1\cap A_2\cap A_3\cap A_4|$$

注意这个公式包含 15 个不同的项，对于$\{A_1, A_2, A_3, A_4\}$的每个非空子集有一项。 ■

练习

1. 在 $A_1\cup A_2$ 中存在多少个元素？如果在 A_1 中存在 12 个元素，在 A_2 中存在 18 个元素，并且
   a) $A_1\cap A_2=\varnothing$　b) $|A_1\cap A_2|=1$　c) $|A_1\cap A_2|=6$　d) $A_1\subseteq A_2$
2. 一个学院有 345 个学生选了微积分课，212 个学生选了离散数学课，188 个学生同时选了微积分和离散数学课。有多少学生选了微积分或离散数学课？
3. 一个调查显示在美国 96%的家庭至少有 1 台电视机，98%的家庭有电话，95%的家庭有电话且至少有 1 台电视机。在美国有百分之几的家庭既没有电话也没有电视机？
4. 一个关于个人计算机的市场报告说 650 000 个拥有计算机的人将在下一年为自己的机器买调制解调器并且 1 250 000 人将至少买 1 个软件包。如果这个报告说 1 450 000 个人将买调制解调器或至少买 1 个软件包，那么有多少人将买调制解调器并且至少买 1 个软件包？
5. 求 $A_1\cup A_2\cup A_3$ 中的元素数，如果每个集合有 100 个元素并且
   a)这些集合是两两不交的。
   b)每对集合中存在 50 个公共元素并且没有元素在所有这 3 个集合里。
   c)每对集合中存在 50 个公共元素并且有 25 个元素在所有这 3 个集合里。
   d)这些集合是相等的。
6. 求 $A_1\cup A_2\cup A_3$ 中的元素数，如果 A_1 中有 100 个元素，A_2 中有 1000 个元素，A_3 中有 10 000 个元素并且
   a) $A_1\subseteq A_2$ 且 $A_2\subseteq A_3$。
   b)这些集合是两两不交的。
   c)在每对集合中存在 2 个公共元素并且没有元素在所有这 3 个集合里。
7. 一个学校有 2504 个计算机科学专业的学生，其中 1876 人选修了 Pascal，999 人选修了 Fortran，345 人选修了 C，876 人选修了 Pascal 和 Fortran，231 人选修了 Fortran 和 C，290 人选修了 Pascal 和 C。如果 189 个学生选了 Fortran，Pascal 和 C，那么 2504 个学生中有多少学生没选这 3 门程序设计语言课的任何一门？
8. 一个关于 270 个大学生的调查显示：64 人喜欢芽甘蓝，94 人喜欢椰菜，58 人喜欢花椰菜，26 人喜欢芽甘蓝和椰菜，28 人喜欢芽甘蓝和花椰菜，22 人喜欢椰菜和花椰菜，14 人喜欢这 3 种蔬菜。270 个学生中有多少人对这 3 种菜都不喜欢？
9. 一个学校有 507，292，312 和 344 个学生分别选了微积分、离散数学、数据结构或程序设计语言课，且有 14 人选了微积分和数据结构课，213 人选了微积分和程序设计语言课，211 人选了离散数学和数据结构课，43 人选了离散数学和程序设计语言课，没有学生同时选微积分和离散数学课，也没有学生同时选数据结构和程序设计语言课。问有多少学生在微积分、离散数学、数据结构或程序设计语言中选了课？
10. 求不超过 100 且不被 5 或 7 整除的正整数个数。
11. 求不超过 100 且是奇数或平方数的正整数个数。
12. 求不超过 1000 且是平方数或立方数的正整数个数。
13. 有多少 8 位二进位串不包含 6 个连续的 0？
*14. 26 个英文字母的排列中有多少个不包含串 fish，rat 或 bird？
15. 在 10 个十进制数字的排列中有多少个以 3 个数字 987 开始，在第 5 和第 6 位包含数字 45，且最后 3 位是 123？
16. 有 4 个集合，每个集合有 100 个元素，每一对集合有 50 个公共元素，每 3 个集合有 25 个公共元素，并

且有 5 个元素在所有的 4 个集合里。问在这 4 个集合的并集中有多少个元素?

17. 有 4 个集合，如果这些集合分别有 50，60，70 和 80 个元素，每一对集合有 5 个公共元素，每 3 个集合有 1 个公共元素，并且没有元素在所有的 4 个集合里。问在这 4 个集合的并集中有多少个元素?
18. 在容斥原理所给出的有关 10 个集合并集元素数的公式中有多少项?
19. 根据容斥原理写出关于 5 个集合并集元素数的显式公式。
20. 有 5 个集合，如果每个集合包含 10 000 个元素，每对集合包含 1000 个公共元素，每 3 个集合包含 100 个公共元素，每 4 个集合包含 10 个公共元素，且这 5 个集合有 1 个公共元素。问在这 5 个集合的并集中有多少个元素?
21. 有 6 个集合，如果知道其中任 3 个集合都是不相交的，根据容斥原理写出关于这 6 个集合并集元素数的显式公式。
*22. 使用数学归纳法证明容斥原理。
23. 设 E_1，E_2 和 E_3 是样本空间 S 的 3 个事件。求一个关于 $E_1 \cup E_2 \cup E_3$ 的概率的公式。
24. 当一个硬币掷 5 次时求头像向下恰好 3 次、第一次和最后一次头像向下、或第二次和第四次头像向上的概率。
25. 从 1 到 100(含 1 和 100 在内)不允许重复地随机取 4 个数，求所有的都是奇数、所有的都被 3 整除或所有的都被 5 整除的概率。
26. 一个样本空间有 4 个事件，如果其中没有 3 个事件同时出现，求关于这 4 个事件的并的概率公式。
27. 一个样本空间有 5 个事件，如果其中没有 4 个事件同时出现，求关于这 5 个事件的并的概率公式。
28. 一个样本空间有 n 个事件，如果其中没有 2 个事件同时出现，求关于这 n 个事件的并的概率公式。
29. 求一个样本空间中 n 个事件的并的概率公式。

6.6 容斥原理的应用

6.6.1 引言

可以使用容斥原理求解许多计数问题。例如，我们可以使用这个原理找出小于某个正整数的素数个数。通过计数从一个有穷集到另一个有穷集的映上函数的个数，能够求解许多问题，而容斥原理就可以用来求出这种函数的个数。也可以使用容斥原理求解著名的帽子认领问题。帽子认领问题是：一个招待随机地将帽子发还存放帽子的人，求没有人取回自己的帽子的概率。

6.6.2 容斥原理的另一种形式

容斥原理有另一种表述形式，它在计数问题中是很有用的。特别地，这种形式可以用于求解在一个集合中的元素数，使得这些元素不具有 n 个性质 P_1，P_2，…，P_n 中的任何一条性质。

设 A_i 是具有性质 P_i 的元素的子集。具有所有这些性质 P_{i_1}，P_{i_2}，…，P_{i_k} 的元素数将记作 $N(P_{i_1}P_{i_2}\cdots P_{i_k})$。用集合的术语写这些等式，有

$$|A_{i_1}\cap A_{i_2}\cap\cdots\cap A_{i_k}| = N(P_{i_1}P_{i_2}\cdots P_{i_k})$$

如果不具有 n 个性质 P_1，P_2，…，P_n 中的任何一条的元素数记作 $N(P'_1P'_2\cdots P'_n)$ 并且集合中的元素数记作 N，那么有

$$N(P'_1P'_2\cdots P'_n) = N - |A_1\cup A_2\cup\cdots\cup A_n|$$

由容斥原理，有

$$N(P'_1P'_2\cdots P'_n) = N - \sum_{1\leqslant i\leqslant n}N(P_i) + \sum_{1\leqslant i<j\leqslant n}N(P_iP_j) - \sum_{1\leqslant i<j<k\leqslant n}N(P_iP_jP_k) + \cdots + (-1)^nN(P_1P_2\cdots P_n)$$

例 1 说明怎样使用容斥原理确定具有约束条件的方程的整数解的个数。

例 1 $x_1+x_2+x_3=11$ 有多少个整数解？其中 x_1，x_2 和 x_3 是非负整数，且 $x_1\leqslant 3$，$x_2\leqslant 4$，$x_3\leqslant 6$。

解 为使用容斥原理，令解的性质 P_1 为 $x_1>3$，性质 P_2 为 $x_2>4$，性质 P_3 为 $x_3>6$。满足不等式 $x_1\leqslant 3$，$x_2\leqslant 4$ 以及 $x_3\leqslant 6$ 的解的个数是

$$\begin{aligned}N(P_1'P_2'P_3') = N - N(P_1) - N(P_2) - N(P_3) + N(P_1P_2)\\ + N(P_1P_3) + N(P_2P_3) - N(P_1P_2P_3)\end{aligned}$$

使用与 4.5 节的例 5 相同的技术，得

$$\begin{aligned}
N &= \text{解的总数} = C(3+11-1,11) = 78\\
N(P_1) &= (\text{具有 } x_1\geqslant 4 \text{ 的解数}) = C(3+7-1,7) = C(9,7) = 36\\
N(P_2) &= (\text{具有 } x_2\geqslant 5 \text{ 的解数}) = C(3+6-1,6) = C(8,6) = 28\\
N(P_3) &= (\text{具有 } x_3\geqslant 7 \text{ 的解数}) = C(3+4-1,4) = C(6,4) = 15\\
N(P_1P_2) &= (\text{具有 } x_1\geqslant 4 \text{ 且 } x_2\geqslant 5 \text{ 的解数}) = C(3+2-1,2) = C(4,2) = 6\\
N(P_1P_3) &= (\text{具有 } x_1\geqslant 4 \text{ 且 } x_3\geqslant 7 \text{ 的解数}) = C(3+0-1,0) = 1\\
N(P_2P_3) &= (\text{具有 } x_2\geqslant 5 \text{ 且 } x_3\geqslant 7 \text{ 的解数}) = 0\\
N(P_1P_2P_3) &= (\text{具有 } x_1\geqslant 4, x_2\geqslant 5 \text{ 且 } x_3\geqslant 7 \text{ 的解数}) = 0
\end{aligned}$$

把这些等式代入关于 $N(P_1'P_2'P_3')$的公式，证明满足 $x_1\leqslant 3$，$x_2\leqslant 4$ 以及 $x^3\leqslant 6$ 的解的个数等于

$$N(P_1'P_2'P_3') = 78-36-28-15+6+1+0-0=6$$

■

6.6.3 埃拉托色尼筛

可以用容斥原理找出不超过一个给定正整数的素数个数。一个合数可以被一个不超过它的平方根的素数整除。因此，为找出不超过 100 的素数个数，首先注意到不超过 100 的合数一定有一个不超过 10 的素因子。由于小于 10 的素数只有 2，3，5 和 7，因此不超过 100 的素数就是这 4 个素数以及那些大于 1 和不超过 100 且不被 2，3，5 或 7 整除的正整数。为应用容斥原理，令 P_1 是一个整数被 2 整除的性质，P_2 是一个整数被 3 整除的性质，P_3 是一个整数被 5 整除的性质，$P4$ 是一个整数被 7 整除的性质。于是，不超过 100 的素数个数是

$$4+N(P_1'P_2'P_3'P_4')$$

由于存在 99 个比 1 大且不超过 100 的正整数，容斥原理显示

$$\begin{aligned}N(P_1'P_2'P_3'P_4') = 99 &- N(P_1) - N(P_2) - N(P_3) - N(P_4)\\ &+ N(P_1P_2) + N(P_1P_3) + N(P_1P_4) + N(P_2P_3) + N(P_2P_4) + N(P_3P_4)\\ &- N(P_1P_2P_3) - N(P_1P_2P_4) - N(P_1P_3P_4) - N(P_2P_3P_4)\\ &+ N(P_1P_2P_3P_4)\end{aligned}$$

不超过 100(且大于 1)并被{2，3，5，7}的子集中的所有素数整除的正整数个数是$\lfloor 100/N\rfloor$，其中 N 是这个子集中的素数之积。(这是由于任意两个素数都没有公因子。)因此

$$\begin{aligned}N(P_1'P_2'P_3'P_4') &= 99 - \left\lfloor\frac{100}{2}\right\rfloor - \left\lfloor\frac{100}{3}\right\rfloor - \left\lfloor\frac{100}{5}\right\rfloor - \left\lfloor\frac{100}{7}\right\rfloor + \left\lfloor\frac{100}{2\cdot 3}\right\rfloor + \left\lfloor\frac{100}{2\cdot 5}\right\rfloor + \left\lfloor\frac{100}{2\cdot 7}\right\rfloor\\
&\quad + \left\lfloor\frac{100}{3\cdot 5}\right\rfloor + \left\lfloor\frac{100}{3\cdot 7}\right\rfloor + \left\lfloor\frac{100}{5\cdot 7}\right\rfloor - \left\lfloor\frac{100}{2\cdot 3\cdot 5}\right\rfloor - \left\lfloor\frac{100}{2\cdot 3\cdot 7}\right\rfloor - \left\lfloor\frac{100}{2\cdot 5\cdot 7}\right\rfloor\\
&\quad - \left\lfloor\frac{100}{3\cdot 5\cdot 7}\right\rfloor + \left\lfloor\frac{100}{2\cdot 3\cdot 5\cdot 7}\right\rfloor\\
&= 99-50-33-20-14+16+10+7+6+4+2-3-2-1-0+0\\
&= 21\end{aligned}$$

于是存在 4＋21＝25 个不超过 100 的素数。

可以用埃拉托色尼筛⊖求不超过一个给定正整数的所有的素数。例如，可用下面的过程找不超过 100 的所有素数。首先，保留 2 而将其余那些被 2 整除的整数删除。因为 3 是保留下来的第一个大于 2 的整数，除 3 之外，删除其余那些被 3 整除的整数。因为 5 是在 3 后面下一个留下来的整数，除 5 之外删除其余那些被 5 整除的整数。下一个留下的整数是 7，因此留下 7，删除其余那些被 7 整除的整数。由于所有不超过 100 的合数被 2，3，5 或 7 整除，那么所有留下来的大于 1 的数是素数。在表 6-2 中，4 个子表显示了每一步删除的整数，在第一个表中得到下划线的是除 2 之外其余被 2 整除的整数，在第二个表中得到下划线的是除 3 之外其余被 3 整除的整数，在第三个表中得到下划线的是除 5 之外其余被 5 整除的整数，在第四个表中得到下划线的是除 7 之外其余被 7 整除的整数。没划线的整数是不超过 100 的素数。

表 6-2　埃拉托色尼筛

除 2 之外其余被 2 整除的整数得到下划线

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

除 3 之外其余被 3 整除的整数得到下划线

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

除 5 之外其余被 5 整除的整数得到下划线

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

除 7 之外其余被 7 整除的整数得到下划线

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

⊖ **埃拉托色尼(Eratosthenes，公元前 276—194)**　只知道埃拉托色尼诞生在塞伦，埃及西部的一个希腊人聚居地，并且在雅典的柏拉图研究院从事研究。托勒密二世皇帝曾邀请埃拉托色尼到亚力山大给他的儿子做家庭教师，后来埃拉托色尼成为著名的亚力山大图书馆馆长。这个图书馆是古代智慧的宝库。埃拉托色尼是一个多才多艺的学者，他的著作涉及数学、地理学、天文学、历史、哲学和文学评论诸方面。除了在数学方面的成就外，他最引人注目的是关于古代历史编年表和地球大小测量的工作。

6.6.4 映上函数的个数

也可以用容斥原理确定从 m 元素集合到 n 元素集合的映上函数的个数。首先考虑例 2。

例 2 从 6 元素集合到 3 元素集合有多少个映上函数?

解 假定在陪域中的元素是 b_1，b_2，b_3。设 P_1，P_2，P_3 分别是 b_1，b_2，b_3 不在函数值域中的性质。注意到一个函数是映上的当且仅当没有性质 P_1，P_2 和 P_3。根据容斥原理得到 6 元素集合到 3 元素集合的映上函数个数是

$$N(P_1'P_2'P_3') = N - [N(P_1) + N(P_2) + N(P_3)] + [N(P_1P_2) + N(P_1P_3) + N(P_2P_3)] - N(P_1P_2P_3)$$

其中 N 是从 6 元素集合到 3 元素集合的函数总数。我们将对等式右边的每一项求值。

由 4.1 节的例 5 得出 $N=3^6$。注意到 $N(P_i)$是值域中不含 b_i 的函数个数。所以，对于定义域中的每个元素的函数值有 2 种选择，从而得到 $N(P_i)=2^6$。此外，这种项有 $C(3, 1)$个。注意到 $N(P_iP_j)$是值域中不含 b_i 和 b_j 的函数个数。所以，对于定义域中的每个元素的函数值只有 1 种选择。从而得到 $N(P_iP_j)=1^6=1$。此外，这种项有 $C(3, 2)$个。还有，注意到 $N(P_1P_2P_3)=0$，因为这个项是值域中不含 b_1，b_2 和 b_3 的函数个数。很显然，没有这样的函数。于是，从 6 元素集合到 3 元素集合的映上函数个数是

$$3^6 - C(3,1)2^6 + C(3,2)1^6 = 729 - 192 + 3 = 540$$ ■

现在叙述从 m 元素集合到 n 元素集合的映上函数个数的一般性结果。这个结果的证明留给读者作为练习。

定理 1 设 m 和 n 是正整数满足 $m \geqslant n$。那么存在

$$n^m - C(n,1)(n-1)^m + C(n,2)(n-2)^m - \cdots + (-1)^{n-1}C(n,n-1) \cdot 1^m$$

个从 m 元素集合到 n 元素集合的映上函数。

注意 从 m 元素集合到 n 元素集合的映上函数的个数等于 $n!\ S_2(m, n)$，其中 $S_2(m, n)$是第二类斯特林数，要了解细节请看[MiRo91]的第 6 章。

下面给出定理 1 的另一个应用的实例。

例 3 把 5 项工作分给 4 个不同的雇员，如果每个雇员至少分配 1 项工作，问有多少种方式?

解 把工作分配看作从 5 个工作集合到 4 个雇员集合的函数。每个雇员至少得到 1 项工作的分配对应于从工作集合到雇员集合的映上函数。因此，由定理 1 存在

$$4^5 - C(4,1)3^5 + C(4,2)2^5 - C(4,3)1^5 = 1024 - 972 + 192 - 4 = 240$$

种方式来分配工作并使得每个雇员至少得到 1 项工作。 ■

6.6.5 错位排列

下面将用容斥原理计数排列 n 个物体并使得没有一个物体在它的初始位置上的方式数。考虑下面的例子。

例 4 帽子认领问题 在一个餐厅里，一个新的雇员寄存 n 个人的帽子时忘记把寄存号放在帽子上。当顾客取回他们的帽子时，这个雇员从剩下的帽子中随机选择发给他们。问没有一个人收到自己的帽子的概率是多少? ■

注意 答案就是重新排列帽子使得没有帽子在它的初始位置上的方式数除以 n 个帽子的排列数 $n!$。在我们找出排列 n 个物体并使得没有一个物体在它的初始位置上的方式数以后再考虑这个例子。

一个错位排列[⊖]是使得没有一个物体在它的初始位置上的排列。为求解例 4 中的问题我们需要确定 n 个物体的错位排列数。

例 5 排列 21453 是 12345 的一个错位排列，因为没有数在它的初始位置上。但是，21543 不是 12345 的错位排列，因为 4 留在它的初始位置上。■

设 D_n 表示 n 个物体的错位排列数。例如，$D_3=2$，因为 123 的错位排列是 231 和 312。我们将使用容斥原理对所有的正整数 n 求 D_n。

定理 2 n 元素集合的错位排列数是

$$D_n=n!\left[1-\frac{1}{1!}+\frac{1}{2!}-\frac{1}{3!}+\cdots+(-1)^n\frac{1}{n!}\right]$$

证 如果排列保持元素 i 不变，就设排列有性质 P_i。错位排列的个数就是对 $i=1, 2, \cdots, n$，没有性质 P_i 的排列数，或

$$D_n = N(P_1'P_2'\cdots P_n')$$

使用容斥原理得到

$$D_n = N-\sum_i N(P_i)+\sum_{i<j}N(P_iP_j)-\sum_{i<j<k}N(P_iP_jP_k)+\cdots+(-1)^nN(P_1P_2\cdots P_n)$$

其中 N 是 n 个元素的排列数。这个等式说明，所有的元素都发生变化的排列数，等于排列的总数减去至少保持 1 个元素不变的排列数，加上至少保持 2 元素不变的排列数，减去至少保持 3 个元素不变的排列数，等等。现在找出在等式右边出现的所有的量。

首先注意到 $N=n!$，因为 N 仅仅就是 n 个元素排列的总数。还有，$N(P_i)=(n-1)!$，这是由乘积法则得到的，因为 $N(P_i)$是保持元素 i 不变的排列数，因而第 i 个位置是确定的，但是其余的每个位置可以放任意元素。类似地，

$$N(P_iP_j) = (n-2)!$$

因为这是保持元素 i 和 j 不变的排列数，但是其余$(n-2)$个元素的位置可以被任意地安排。一般说来有

$$N(P_{i_1}P_{i_2}\cdots P_{i_m}) = (n-m)!$$

因为这是保持元素 i_1，i_2，…，i_m 不变的排列数，但是其他$(n-m)$个元素的位置可以被任意安排。由于存在 $C(n, m)$种方式从 n 个元素中选择 m 个，从而有

$$\sum_{1\leqslant i\leqslant n}N(P_i) = C(n,1)(n-1)!$$

$$\sum_{1\leqslant i<j\leqslant n}N(P_iP_j) = C(n,2)(n-2)!$$

一般地有

$$\sum_{1\leqslant i_1<i_2<\cdots<i_m\leqslant n}N(P_{i_1}P_{i_2}\cdots P_{i_m}) = C(n,m)(n-m)!$$

所以，把这些等式代入关于 D_n 的公式得

$$\begin{aligned}D_n &=n!-C(n,1)(n-1)!+C(n,2)(n-2)!-\cdots+(-1)^nC(n,n)(n-n)!\\&=n!-\frac{n!}{1!(n-1)!}(n-1)!+\frac{n!}{2!(n-2)!}(n-2)!-\cdots+(-1)^n\frac{n!}{n!0!}0!\end{aligned}$$

⊖ **历史注记** 在一个古老的法国纸牌相遇(匹配)游戏中，一套 52 张牌摆成一行。摆放第二套牌使其中每张牌放在第一套牌的某一张的顶部。通过统计在两套牌中匹配的牌数来确定得分。在 1708 年，皮埃尔·雷蒙德·蒙特莫特(1678—1719)提出了“相遇”问题：在相遇游戏中没有匹配发生的概率是多少？蒙特莫特问题的解是随机选择 52 个物体的排列恰为错位排列的概率，即 $D_{52}/52!$，正如我们将看到的，这个概率近似为 $1/e$。

化简这个表达式得

$$D_n = n!\left[1-\frac{1}{1!}+\frac{1}{2!}-\cdots+(-1)^n\frac{1}{n!}\right]$$ □

现在对于给定的正整数 n 求 D_n 就简单了。例如，使用定理 2 得

$$D_3 = 3!\left[1-\frac{1}{1!}+\frac{1}{2!}-\frac{1}{3!}\right]=6\left(1-1+\frac{1}{2}-\frac{1}{6}\right)=2$$

正如我们前面所看到的。

现在可以给出例 4 中的问题的解。

解 没有一个人收到自己的帽子的概率是 $D_n/n!$。由定理 2，这个概率是

$$\frac{D_n}{n!}=1-\frac{1}{1!}+\frac{1}{2!}-\cdots+(-1)^n\frac{1}{n!}$$

对于 $2\leqslant n\leqslant 7$，这个概率的值在表 6-3 中给出。

表 6-3 错位排列的概率

n	2	3	4	5	6	7
$D_n/n!$	0.500 00	0.333 33	0.375 00	0.366 67	0.368 06	0.367 86

使用微积分的方法可以证明

$$e^{-1}=1-\frac{1}{1!}+\frac{1}{2!}-\cdots+(-1)^n\frac{1}{n!}+\cdots\approx 0.368$$

因为这是一个项趋向于 0 的交错级数，当 n 无限增长时，没有一个人取回自己的帽子的概率趋于 $e^{-1}\approx 0.368$。事实上，可以证明这个概率与 e^{-1} 的差在 $1/(n+1)!$ 之内。 ■

练习

1. 假设 1 蒲式耳 100 个苹果中 20 个有虫，15 个有擦伤。只有没虫也没擦伤的苹果才可以卖。如果 10 个擦伤的苹果有虫，那么 100 个苹果中有多少个可以卖？
2. 1000 个人申请喜马拉雅登山旅游，450 人有高山病，622 人不是处在很好的状态，30 人有过敏症。一个申请人当且仅当没有高山病，是在一个良好的状态下，并且没有过敏症，才算合格。如果 111 个申请人有高山病且不是在良好状态，14 人有高山病和过敏症，18 人不是在良好的状态并且有过敏症。9 个人有高山病并且不是在良好状态和有过敏症。那么有多少申请人合格？
3. 方程 $x_1+x_2+x_3=13$ 有多少个解？其中 x_1，x_2，x_3 是小于 6 的非负整数。
4. 求方程 $x_1+x_2+x_3+x_4=17$ 的解的个数，其中 $x_i(i=1, 2, 3, 4)$是非负整数，满足条件 $x_1\leqslant 3$，$x_2\leqslant 4$，$x_3\leqslant 5$ 且 $x_4\leqslant 8$。
5. 使用容斥原理求小于 200 的素数的个数。
6. 一个整数叫做无平方因子的如果它不被一个大于 1 的正整数的平方整除。求小于 100 的无平方的正整数个数。
7. 有多少小于 10 000 的正整数不是一个整数的 2 次或更高次幂？
8. 从 7 元素集合到 5 元素集合有多少个映上函数？
9. 有多少种方式把 6 个不同的玩具分给 3 个不同的孩子并使得每个孩子至少得到 1 个玩具？
10. 8 个不同的球放入 3 个不同的罐子，如果每个罐子至少有 1 个球，有多少种放法？
11. 有多少种方式把 7 项不同的工作分给 4 个不同的雇员，使得每个雇员至少得到 1 项工作，并且把最困难的工作分给最好的雇员？
12. 列出{1，2，3，4}的所有的错位排列。
13. 一个 7 元素集合有多少个错位排列？

14. 如果寄存帽子的人随机发回帽子，10 个人中没有一个人得到他自己帽子的概率是多少？
15. 一个把信放入信袋的机器发生了故障并且随机把信放入信袋中。在一组 100 封信中发生下面事件的概率是多少？
   a)没有信放对了信袋。
   b)恰好 1 封信放对了信袋。
   c)恰好 98 封信放对了信袋。
   d)恰好 99 封信放对了信袋。
   e)所有的信都放对了信袋。
16. 在同一个教室为一组 n 个学生分配两个班的座位。如果没有学生在两个班分到同一个座位，有多少种方式？
*17. 有多少种方式安排数字 0，1，2，3，4，5，6，7，8，9 使得没有偶数在它的初始位置上？
*18. 设 D_n 表示 n 个物体的错位排列数，用组合论证证明序列$\{D_n\}$满足递推关系

$$D_n=(n-1)(D_{n-1}+D_{n-2}) \quad 对于\ n\geqslant 2$$

(提示：对于错位排列中的第 1 个元素 k 有 $n-1$ 种选择。单独考虑以 k 开始的错位排列，它的第 k 位可以是 1 也可以不是 1。)
*19. 使用练习 18 证明

$$D_n=nD_{n-1}+(-1)^n \quad 对于\ n\geqslant 1$$

20. 使用练习 19 求关于 D_n 的显式公式。
21. 对哪些正整数 n，错位排列数 D_n 是偶数？
22. 假设 p 和 q 是不同的素数。使用容斥原理求 $\phi(pq)$，即不超过 pq 且与 pq 互素的整数个数。
*23. 当 n 的素因子分解式是

$$n=p_1^{a_1}p_2^{a_2}\cdots p_m^{a_m}$$

时，使用容斥原理推导一个关于 $\phi(n)$ 的公式。
*24. 证明如果 n 是正整数，那么

$$n!=C(n,0)D_n+C(n,1)D_{n-1}+\cdots+C(n,n-1)D_1+C(n,n)D_0$$

其中 D_k 是 k 个物体的错位排列数。
25. 有多少个{1，2，3，4，5，6}的错位排列的前 3 位是取某种次序的整数 1，2，3？
26. 有多少个{1，2，3，4，5，6}的错位排列的后 3 位是取某种次序的整数 1，2，3？
27. 证明定理 1。

关键术语和结果

术语

递推关系：一个公式，它把序列除了某些初始项以外的项，表示成这个序列前面的一个或若干个项的函数。

递推关系的初始条件：满足递推关系的序列在该关系起作用之前的某些项的值。

常系数线性齐次递推关系：一个递推关系，除了初始项之外，它把序列的项表示成前面项的线性组合。

常系数线性齐次递推关系的特征根：与常系数线性齐次递推关系相关的多项式的根。

常系数线性非齐次递推关系：一个递推关系，除了初始项之外，它把序列的项表示成前面项的线性组合加上一个仅仅依赖于序标的不恒为 0 的函数。

分治算法：求解问题的一种算法，求解中递归地把问题划分成固定数目的较小的同种类型的问题。

序列的生成函数：用序列的第 n 项作为 x^n 的系数的形式幂级数。

埃拉托色尼筛：找出小于一个给定正整数的素数的过程。

错位排列：使得没有物体在它的初始位置上的排列。

结果

两个有穷集合并集的元素数公式：

$$|A \cup B| = |A| + |B| - |A \cap B|$$

三个有穷集合并集的元素数公式：

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$$

容斥原理：

$$|A_1 \cup A_2 \cup \cdots \cup A_n| = \sum_{1 \leqslant i \leqslant n} |A_i| - \sum_{1 \leqslant i < j \leqslant n} |A_i \cap A_j| + \sum_{1 \leqslant i < j < k \leqslant n} |A_i \cap A_j \cap A_k| - \cdots + (-1)^{n+1} |A_1 \cap A_2 \cap \cdots \cap A_n|$$

从 m 元素集合到 n 元素集合的映上函数个数：

$$n^m - C(n,1)(n-1)^m + C(n,2)(n-2)^m - \cdots + (-1)^{n-1} C(n,n-1) \cdot 1^m$$

n 个物体的错位排列数：

$$D_n = n!\left[1 - \frac{1}{1!} + \frac{1}{2!} - \cdots + (-1)^n \frac{1}{n!}\right]$$

复习题

1. a)什么是递推关系？
   b)如果在一个获利 9%的账上储蓄 1 000 000 美元，求与 n 年后账上钱数有关的递推关系。
2. 解释怎样用斐波那契数求解关于兔子的斐波那契问题。
3. a)找出与求解汉诺塔难题的所需步数有关的递推关系。
   b)显示怎样使用迭代来求解这个递推关系。
4. a)解释怎样找一个与不包含两个连续的 1 的 n 位二进位串个数有关的递推关系。
   b)描述另一个计数问题使得它的解满足同一个递推关系。
5. 定义一个 k 阶的线性齐次递推关系。
6. a)解释怎样求解二阶线性齐次递推关系。
   b)如果 $a_0=3$，$a_1=15$，对于 $n \geqslant 2$，求解递推关系 $a_n=13a_{n-1}-22a_{n-2}$。
   c)如果 $a_0=3$，$a_1=35$，对于 $n \geqslant 2$，求解递推关系 $a_n=14a_{n-1}-49a_{n-2}$。
7. a)如果 $f(n)$满足分治递推关系 $f(n)=af(n/b)+g(n)$，这里 b 整除正整数 n，解释怎样求 $f(b^k)$，其中 k 是正整数。
   b)如果 $f(n)=3f(n/4)+5n/4$ 且 $f(1)=7$，求 $f(256)$。
8. a)对于用二分搜索在表中找一个数所用的比较次数推导一个分治的递推关系。
   b)从你在(a)给出的分治递推关系使用 6.3 节中的定理 1，对于二分搜索所用的比较次数给出一个大 O 估计。
9. a)给出一个关于 3 个集合并集元素数的公式。
   b)解释为什么这个公式是有效的。
   c)解释怎样使用(a)的公式求不超过 1000 且能被 6，10 或 15 整除的正整数个数。
   d)解释怎样使用(a)的公式求方程 $x_1+x_2+x_3+x_4=22$ 的非负整数解的个数，其中 $x_1<8$，$x_2<6$，$x_3<5$。
10. a)给出一个关于 4 个集合并集元素数的公式。解释为什么它是有效的。

b)假设 A_1，A_2，A_3 和 A_4 每个集合含 25 个元素，其中任 2 个集合的交含 5 个元素，任 3 个集合的交含 2 个元素，所有 4 个集合含 1 个公共元素。问在这 4 个集合的并集中有多少个元素？

11. a)叙述容斥原理。

 b)概述这个原理的证明。

12. 解释怎样使用容斥原理计数从 m 元素集合到 n 元素集合的映上函数的个数。

13. a)怎样计数把 m 项工作分给 n 个雇员并使得每个雇员至少得到一项工作的方案数？

 b)把 7 项工作分给 3 个雇员并使得每个雇员至少得到一项工作有多少种方案？

14. 解释怎样使用容斥原理计数不超过正整数 n 的素数个数。

15. a)定义一个错位排列。

 b)一个寄存帽子的人给 n 个人发还帽子，使得没有人得到自己帽子的方式的计数为什么和 n 个物体的错位排列数一样？

 c)解释怎样计数 n 个物体的错位排列数。

补充练习

1. 一个 10 人小组开始一系列的通信活动，每个人把这封信寄给另外 4 个人。每个收到信的人再把这封信寄给另外的 4 个人。

 a)如果没有人收到的信多于 1 封，求与这个通信活动的第 n 步寄出信数有关的递推关系。

 b)在(a)中递推关系的初始条件是什么？

 c)在这个通信活动的第 n 步寄出了多少封信？

2. 一个核反应堆产生了 18 克放射性同位素。每小时放射性同位素衰变 1%。

 a)对 n 小时后留下的同位素量建立一个递推关系。

 b)对于(a)的递推关系，初始条件是什么？

 c)求解这个递推关系。

3. 美国政府每小时印 1 美元纸币超过 10 000 张，5 美元纸币超过 4000 张，10 美元纸币超过 3000 张，20 美元纸币超过 2500 张，50 美元纸币超过 1000 张，100 美元纸币与前一小时的张数一样。在初始时刻每种钱币有 1000 张。

 a)建立一个关于第 n 小时总钱数的递推关系。

 b)(a)中递推关系的初始条件是什么？

 c)求解这个第 n 小时总钱数的递推关系。

 d)建立一个关于前 n 小时总钱数的递推关系。

 e)求解这个关于前 n 小时总钱数的递推关系。

4. 每个前一小时已经存在的细菌在每小时都分裂出 2 个新的细菌，并且所有的细菌只有 2 小时的寿命。假设这群细菌开始时有 100 个新细菌。

 a)建立关于 n 小时后存在细菌数目的递推关系。

 b)这个递推关系的解是什么？

 c)什么时候这群细菌的个数将超过 100 万个？

5. 使用两个不同的信号在通信信道发送信息。传送一个信号需要 2 微秒，传送另一个信号要 3 微秒。一个信息的每个信号后紧跟着下一个信号。

 a)求与在 n 微秒中可以发送的不同信号数有关的递推关系。

 b)对于(a)的递推关系，初始条件是什么？

 c)在 12 微秒内可以发送多少个不同的信息？

6. 一个小邮局只有 4 分、6 分和 10 分邮票。如果考虑邮票使用的次序，求与这些邮票构成 n 分邮资的方式数有关的递推关系。这个递推关系的初始条件是什么？

7. 使用在练习 6 描述的规则，有多少种方式构成下述邮资？

 a)12 分　b)14 分　c)18 分　d)22 分

8. 求具有 $a_0=1$ 和 $b_0=2$ 的联立方程组

$a_n=a_{n-1}+b_{n-1}$

$b_n = a_{n-1} - b_{n-1}$

9. 如果 $a_0=1$ 且 $a_1=2$，求解递推关系 $a_n = a_{n-1}^2 / a_{n-2}$。[提示：两边取对数得到关于序列 $\log a_n$ 的递推关系，$n=0, 1, 2, \cdots$。]

*10. 如果 $a_0=2$ 和 $a_1=2$，求解递推关系 $a_n = a_{n-1}^3 a_{n-2}^2$。(见练习 9 的提示。)

11. 如果 $a_0=2$，$a_1=4$ 和 $a_2=8$，求解递推关系 $a_n = 3a_{n-1} - 3a_{n-2} + a_{n-3} + 1$。

12. 如果 $a_0=2$，$a_1=2$ 和 $a_2=4$，求解递推关系 $a_n = 3a_{n-1} - 3a_{n-2} + a_{n-3}$。

*13. 假设在 6.1 节的例 4 中，一对兔子在繁殖 2 次以后就离开这个岛。求与第 n 个月中的岛上兔子对数有关的递推关系。

14. 当 n 可被 5 整除时求解递推关系 $f(n) = 3f(n/5) + 2n^4$，其中 $n=5^k$，k 是正整数，$f(1)=1$。

15. 如果 f 是增函数，估计练习 14 中 f 的大小。

16. 找出与下述算法所使用比较次数有关的递推关系：通过把 n 个数的序列递归地划分成两个子序列找出最大和第二大的元素，在每一步划分时要求这两个子序列项数相等或一个子序列比另一个子序列多一项。当子序列达到 2 项时停止。

17. 估计练习 16 描述的算法所使用的比较次数。

设 $\{a_n\}$ 是实数序列。这个序列的前向差分定义如下：第一个前向差分是 $\Delta a_n = a_{n+1} - a_n$；第 $k+1$ 个前向差分 $\Delta^{k+1} a_n$ 是通过 $\Delta^{k+1} a_n = \Delta^k a_{n+1} - \Delta^k a_n$ 由 $\Delta^k a_n$ 得到的。

18. 求 $\Delta^k a_n$，其中

a) $a_n = 3$　　b) $a_n = 4n+7$　　c) $a_n = n^2+n+1$

19. 设 $a_n = 3n^3+n+2$，求 $\Delta^k a_n$，其中 k 等于

a) 2　　b) 3　　c) 4

*20. 假设 $a_n = P(n)$，其中 P 是 d 次多项式。证明对所有的非负整数 n，$\Delta^{d+1} a_n = 0$。

21. 令 $\{a_n\}$ 和 $\{b_n\}$ 是实数序列。证明

$$\Delta(a_n b_n) = a_{n+1}(\Delta b_n) + b_n(\Delta a_n)$$

22. 证明如果 $F(x)$ 和 $G(x)$ 分别是序列 $\{a_k\}$ 和 $\{b_k\}$ 的生成函数，且 c 和 d 是实数，那么 $(cF+dG)(x)$ 是 $\{ca_k + db_k\}$ 的生成函数。

23. (需要微积分)这个练习说明了怎样使用生成函数求解递推关系 $(n+1)a_{n+1} = a_n + (1/n!)$，$n \geqslant 0$，初始条件 $a_0 = 1$。

a) 设 $G(x)$ 是关于 $\{a_n\}$ 的生成函数。证明 $G'(x) = G(x) + e^x$ 且 $G(0)=1$。

b) 由(a)证明 $(e^{-x}G(x))' = 1$，且断定 $G(x) = xe^x + e^x$。

c) 使用(b)找出关于 a_n 的封闭公式。

24. 假设在离散数学班的第一次考试中 14 个学生得 A，第二次考试中 18 个得 A。如果 22 个学生在第一或第二次考试得 A，有多少学生两次考试都得 A?

25. 在蒙默思郡(英国威尔士郡原郡名)323 个农场至少有马、牛或羊其中的 1 种。如果 224 个农场有马，85 个有牛，57 个有羊，18 个农场 3 种家畜全有，那么有多少个农场恰好有这 3 种家畜中的 2 种?

26. 查询某学院关于学生记录的数据库得到下述数据：学院有 2175 个学生，其中 1675 个不是一年级学生，1074 个学生选了微积分，444 个学生选了离散数学，607 个不是一年级学生且选了微积分，350 个学生选了微积分和离散数学，201 个不是一年级学生且选了离散数学，143 个不是一年级学生并且选了微积分和离散数学。所有这些对查询的回答都是正确的吗?

27. 某大学数学学院的学生可以选择下述一个或多个方向作为主修方向：应用数学(AM)、纯粹数学(PM)、运筹学(OR)及计算机科学(CS)。如果包括同时主修在内，主修 AM 的有 23 个学生；主修 PM 的有 17 个学生；主修 OR 的 44 个；主修 CS 的 63 个；主修 AM 与 PM 的 5 个；主修 AM 与 CS 的 8 个；主修 AM 与 OR 的 4 个；主修 PM 与 CS 的 6 个；主修 PM 与 OR 的 5 个；主修 OR 与 CS 的 14 个；主修 PM，OR 与 CS 的 2 个；主修 AM，OR 与 CS 的 2 个；主修 PM，AM 与 OR 的 1 个；主修 PM，AM 与

CS 的 1 个；还有 1 个主修所有 4 个方向。问这个学院有多少学生？

28. 当使用容斥原理表示 7 个集合的并集中的元素数时，如果其中没有 6 个或更多的集合含有公共元素，那么需要多少项？
29. 方程 $x_1+x_2+x_3=20$，$2<x_1<6$，$6<x_2<10$，$0<x_3<5$ 有多少个正整数解？
30. 有多少小于 1 000 000 的正整数
    a)能够被 2，3 或 5 整除？ b)不被 7，11 或 13 整除？ c)被 3 整除但不被 7 整除？
31. 有多少小于 200 的正整数是
    a)整数的 2 次或更高次幂？ b)整数的 2 次或更高次幂，或是素数？
    c)不被一个大于 1 的整数的平方整除？ d)不被一个大于 1 的整数的立方整除？
    e)不被 3 个或更多的素数整除？
*32. 把 6 个不同的工作分给 3 个不同的雇员，如果最难的工作分给最有经验的雇员并且最容易的工作分给最缺乏经验的雇员，那么有多少种分法？
33. 由寄存帽子的人随机发还给 n 个人帽子，那么恰好一个人拿到自己帽子的概率是多少？
34. 有多少个 6 位二进位串不包含 4 个连续的 1？
35. 一个 6 位二进位串包含至少 4 个 1 的概率是多少？

计算机题目

按下述给定的输入和输出写出程序。

1. 给定正整数 n，列出汉诺塔难题从一根柱子到另一根柱子依照游戏规则移动 n 个盘子需要的所有移动。
2. 给定正整数 n 和整数 k，$1\leqslant k\leqslant n$，列出富雷姆-斯图尔特算法(见 6.1 节练习 54 前面的说明)依照游戏规则用 4 根柱子从一根柱子到另一根柱子移动 n 个盘子需要的所有移动。
3. 给定正整数 n，列出不包含连续 2 个 0 的所有的 n 位二进制序列。
4. 给定正整数 n，写出在 $n+1$ 个变量的乘积中加括号的所有方式。
5. 给定递推关系 $a_n=c_1a_{n-1}+c_2a_{n-2}$ 和正整数 k，其中 c_1 和 c_2 是实数，初始条件为 $a_0=C_0$ 及 $a_1=C_1$，使用迭代求 a_k。
6. 给定递推关系 $a_n=c_1a_{n-1}+c_2a_{n-2}$ 和初始条件 $a_0=C_0$ 及 $a_1=C_1$，确定唯一的解。
7. 给定形如 $f(n)=af(n/b)+c$ 的递推关系，其中 a 是实数，b 是正整数，c 是实数，k 是正整数，使用迭代求 $f(b^k)$。
8. 给定 3 个集合的交集中的元素个数，每两个集合的交集中的元素个数，和每个集合中的元素个数，求其并集中的元素个数。
9. 给定正整数 n，求在 n 个集合的并集中元素个数的公式。
10. 给定正整数 m 和 n，求从 m 元素集合到 n 元素集合的映上函数个数。
11. 给定正整数 n，列出集合$\{1, 2, 3, \cdots, n\}$的所有错位排列。

计算和研究

使用一个计算程序或你已完成的程序做下面的练习。

1. 求 f_{100}，f_{500} 和 f_{1000} 的精确值，其中 f_n 是斐波那契数。
2. 求比 1 000 000 大、比 1 000 000 000 大和比 1 000 000 000 000 大的最小的斐波那契数。
3. 求尽可能多的同为素数的斐波那契数，目前还不知道是否存在无限多个这样的数。
4. 写出求解 10 个盘子的汉诺塔难题所需要的所有的移动。
5. 按照雷夫难题的规则，用 4 根柱子从一根柱到另一根柱移动 20 个盘子，写出使用弗雷姆-斯图尔特算法需要的所有移动。
6. 通过下面的方法验证求解 n 个盘子的雷夫难题的弗雷姆猜想：对于尽可能多的整数 n，证明这个难题不

可能使用比具有最优选择 k 的弗雷姆-斯图尔特算法还要少的移动来求解。

7. 计算对于各种整数 n，包括 16，64，256 和 1024，使用在 6.3 节描述的快速乘法和整数相乘的标准算法(2.5 节算法 3)做两个 n 位整数相乘所需要的运算次数。
8. 计算对于各种整数 n，包括 4，16，64 和 128，使用在 6.3 节描述的快速矩阵乘法和矩阵相乘的标准算法(2.7 节算法 1)做两个 $n\times n$ 矩阵相乘所需要的运算次数。
9. 使用埃拉托色尼筛求不超过 1000 的所有素数。
10. 使用在 6.6 节描述的求不超过 100 的素数个数的方法求不超过 10 000 的素数个数。
11. 列出{1，2，3，4，5，6，7，8}的所有的错位排列。
12. 对所有不超过 20 的正整数 n 计算 n 个物体的一个排列是错位排列的概率，并确定这些概率逼近 e. 的速度。

写作题目

用本书以外的资料，按下列要求写成短文。

1. 找出斐波那契发表的关于兔子数模型难题的原始材料。讨论斐波那契提出的这个问题和其他问题，并且给出关于斐波那契本人的某些信息。
2. 解释斐波那契数怎样在其他应用中出现，如叶序、植物叶片排列的研究、镜子反射的研究等。
3. 描述汉诺塔难题的各种不同的变形问题，包括多于 3 个柱子的(包括课本和练习中讨论的雷夫难题在内)，盘子移动受限制的，以及允许有同样大小盘子的。关于求解每种变形问题所要求的移动次数有什么已知的结论？
4. 尽可能多地讨论出现卡塔兰数的不同的问题。
5. 描述派尔克发现的与一次谎话搜索相关的乌拉姆问题(见 6.3 节的练习 28)的解。
6. 讨论派尔克发现的与多次谎话搜索相关的乌拉姆问题的变种(见 6.3 节的练习 28)，关于这个问题你还知道什么？
7. 讨论求平面上点集的凸包的分治算法。
8. 查询*幸运数*(lucky number)的定义。解释怎样使用类似于埃拉托色尼筛的技术找出幸运数。求所有小于 1000 的幸运数。
9. 描述在数论中使用的筛法。使用这种方法已经得到了哪些结果？
10. 查询古代法国纸牌相遇游戏的规则。描述这些规则并且描述皮埃尔·雷蒙德·蒙特莫特关于"相遇问题"的论文。
11. 描述怎样使用指数生成函数求解各种计数问题。
12. 描述计数的 Polyá 理论和可使用这个理论求解的计数问题的种类。
13. *管家问题*是求安排 n 对夫妇围圆桌就座的方法数，使得就座时男女相间并且没有丈夫和妻子相邻。解释怎样用卢卡斯(E. Lucas)方法求解这个问题。
14. 解释怎样使用*棋盘多项式*(rook polynomial)求解计数问题。

第7章 关　　系

在许多情况下集合的元素之间都存在某种关系。每天我们都要涉及各种关系，例如在一个商行和它的电话号码之间的关系，雇员与其工资之间的关系，一个人与一个亲属之间的关系，等等。在数学中我们研究的关系，如一个正整数与被它除的一个正整数、一个整数与和它模5同余的一个整数、一个实数与一个比它大的实数之间的关系，等等。在计算机科学中常常出现的关系，如一个程序与它所使用的一个变量、一种计算机语言与这个语言的一个有效语句之间的关系等。

集合的元素之间的关系被表示成一种结构，这种结构叫做关系。可以用关系来求解问题，例如确定在一个网络中的哪两个城市之间开通航线，为一个复杂课题的不同阶段的工作找一个可行的执行次序，或者产生一个有用的方式以便在计算机数据库中存储信息。

7.1 关系及其性质

7.1.1 引言

可以用两个相关元素构成的有序对来表达两个集合的元素之间的关系，这是一种最直接的方式。为此，有序对的集合就叫做二元关系。在这一节，我们引入用于描述二元关系的基本术语。在这一章的后面，我们将使用关系来求解涉及通信网络、项目调度以及识别集合中具有共同性质的元素等问题。

定义1　设A和B是集合，一个从A到B的二元关系是$A\times B$的子集。

换句话说，一个从A到B的二元关系是有序对的集合R，其中每个有序对的第一个元素取自A并且第二个元素取自B。我们使用记号aRb表示$(a, b)\in R$，$a\not R b$表示$(a, b)\notin R$。当(a, b)属于R时称a与b有关系R。

二元关系表示了两个集合的元素之间的关系。在本章的后面我们将引入n元关系，它表示了在三个以上集合中元素之间的关系。当不发生混淆时我们将省去二元这个词。

例1～例3说明了关系的概念。

例1　设A是你们学校的学生的集合，B是课程的集合。令R是由(a, b)对构成的关系，其中a是选修课程b的学生。例如，如果Jason Goodfriend和Deborah Sherman选修CS518，即离散数学，有序对(Jason Goodfriend，CS518)和(Deborah Sherman，CS518)属于R。如果Jason Goodfriend也选修CS510，即数据结构，那么有序对(Jason Goodfriend，CS510)也属于R。但是，如果Deborah Sherman没有选修CS510，那么有序对(Deborah Sherman，CS510)不在R中。

注意如果一个学生目前没有选修任何课程，那么在R中没有以这个学生为第一元素的对。类似地，如果一门课程目前没有开设，那么在R中也没有以这门课程作为第二个元素的对。■

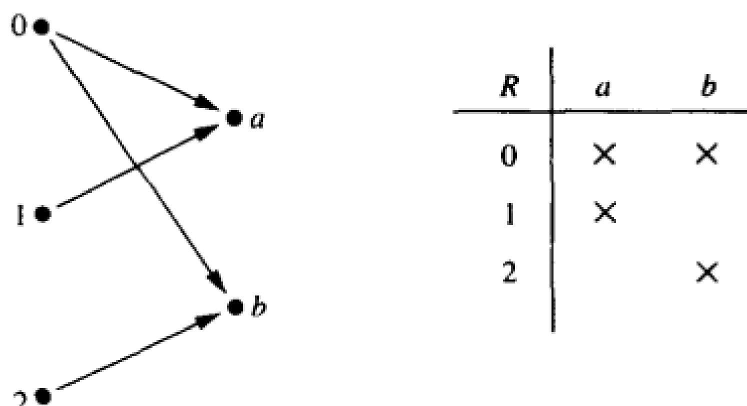
例2　设A是所有城市的集合，B是美国的50个州的集合。如下定义关系R：如果城市a是在州b，则(a, b)属于R。例如，(Boulder，科罗拉多州)，(Bangor，缅因州)，(Ann Arbor，密执安州)，(Cupertino，加利福尼亚州)和(Red Bank，新泽西州)是在R中。■

例3　设$A=\{0, 1, 2\}$，$B=\{a, b\}$，那么$\{(0, a), (0, b), (1, a), (2, b)\}$是从$A$到$B$的关系。这意味着，比如说有$0Ra$，但没有$1Rb$，关系可以用图来表示，如图7-1所示，用箭头

来表示有序对。另一种表示关系的方式就是用一张表，这也在图 7-1 给出。在 7.3 节我们将更详细地讨论关系的表示。 ■

7.1.2 函数作为关系

一个从集合 A 到集合 B 的函数 f(如 1.8 节的定义)对于 A 中的每个元素都指定 B 中一个唯一的元素。f 的图示是使得 $b=f(a)$ 的有序对 (a, b) 的集合。由于 f 的图示是 $A\times B$ 的子集，它就是一个从 A 到 B 的关系。此外，函数的图示有下述性质：A 的每个元素恰好是图中一个有序对的第一元素。



R	a	b
0	×	×
1	×	
2		×

图 7-1　在例 3 的关系 R 中的有序对

相反，如果 R 是从 A 到 B 的关系，并且使得 A 中的每个元素是 R 中恰好一个有序对的第一元素，那么 R 的图示就可以定义一个函数。只要对 A 的每个元素指定唯一的元素 $b\in B$ 使得 $(a, b)\in R$ 就可以做到。

可以用关系表达在集合 A 和集合 B 之间的一对多的相关性，其中 A 的一个元素可以与 B 中多个元素有关。函数表示了这样一种关系，对于 A 中的每个元素恰好只有一个 B 中的元素与之相关。

关系是函数的推广，可以用关系表示集合之间的更为广泛的联系。

7.1.3 集合的关系

集合 A 到它自身的关系是特别令人感兴趣的。

定义 2　集合 A 的关系是从 A 到 A 的关系。

换句话说，集合 A 的关系是 $A\times A$ 的子集。

例 4　设 A 是集合 $\{1, 2, 3, 4\}$，A 上的关系 $R=\{(a, b)\mid a$ 整除 $b\}$ 中有哪些有序对？

解　(a, b) 在 R 中，当且仅当 a 和 b 是不超过 4 的正整数且 a 整除 b，我们看到

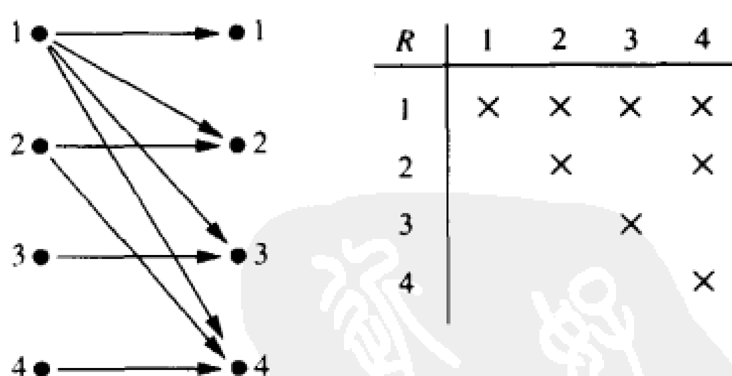
$$R=\{(1,1),(1,2),(1,3),(1,4),(2,2),(2,4),(3,3),(4,4)\}$$

在这个关系中的有序对的图和表的表示都在图 7-2 中给出。 ■

例 5 给出了某些整数集合上的关系的实例。

例 5　考虑下面这些整数集合上的关系：

$$R_1=\{(a,b)\mid a\leqslant b\}$$
$$R_2=\{(a,b)\mid a>b\}$$
$$R_3=\{(a,b)\mid a=b \text{ 或 } a=-b\}$$
$$R_4=\{(a,b)\mid a=b\}$$
$$R_5=\{(a,b)\mid a=b+1\}$$
$$R_6=\{(a,b)\mid a+b\leqslant 3\}$$



R	1	2	3	4
1	×	×	×	×
2		×		×
3			×	
4				×

图 7-2　例 4 中关系 R 的有序对

其中哪些关系包含了有序对 $(1, 1)$、$(1, 2)$、$(2, 1)$、$(1, -1)$ 以及 $(2, 2)$？

注意　和例 1～例 4 的关系不同，这些是无穷集合上的关系。

解　有序对 $(1, 1)$ 在 R_1，R_3，R_4 和 R_6 中；有序对 $(1, 2)$ 在 R_1 和 R_6 中；有序对 $(2, 1)$ 在 R_2，R_5 和 R_6 中；$(1, -1)$ 在 R_2，R_3 和 R_6 中；最后，有序对 $(2, 2)$ 在 R_1，R_3 和 R_4 中。 ■

不难确定有穷集上的关系个数，因为集合 A 上的关系仅仅是 $A\times A$ 的子集。

例6 n元素集合有多少个关系？

解 集合A上的关系是$A\times A$的子集。因为当A是n元素集合时$A\times A$有n^2个元素，并且m个元素的集合有2^m个子集，故$A\times A$的子集有2^{n^2}个。于是n元素集合有2^{n^2}个关系。例如，在集合$\{a, b, c\}$上存在$2^{3^2}=2^9=512$个关系。■

7.1.4 关系的性质

有若干个用于把集合上的关系分类的性质。这里我们只介绍其中最重要的性质。

在某些关系中一个元素总是与自己相关。例如，设R是所有的人的集合上的关系，若x和y有相同的母亲和相同的父亲，那么(x, y)属于R。于是对于每个人x，有xRx。

定义3 如果对每个元素$a\in A$有$(a, a)\in R$，那么集合A上的关系R叫做自反的。

注意 我们看到如果$\forall a((a, a)\in R)$，那么集合上的关系是自反的，这里的论域是A中所有元素的集合。

我们看到如果A的每个元素都关系到它自己，A上的关系就是自反的。例7至例9说明了自反关系的概念。

例7 考虑下面$\{1, 2, 3, 4\}$上的关系

$R_1=\{(1,1),(1,2),(2,1),(2,2),(3,4),(4,1),(4,4)\}$

$R_2=\{(1,1),(1,2),(2,1)\}$

$R_3=\{(1,1),(1,2),(1,4),(2,1),(2,2),(3,3),(4,1),(4,4)\}$

$R_4=\{(2,1),(3,1),(3,2),(4,1),(4,2),(4,3)\}$

$R_5=\{(1,1),(1,2),(1,3),(1,4),(2,2),(2,3),(2,4),(3,3),(3,4),(4,4)\}$

$R_6=\{(3,4)\}$

其中哪些是自反的？

解 关系R_3和R_5是自反的，因为它们都包含了所有形如(a, a)的对，即$(1, 1)$，$(2, 2)$，$(3, 3)$和$(4, 4)$。其他的关系不是自反的，因为它们不包含所有这些有序对。特别地，R_1，R_2，R_4和R_6不是自反的，因为$(3, 3)$都不在这些关系里。■

例8 例5中哪些关系是自反的？

解 这个例子中的自反关系是R_1(因为对每个整数a有，$a\leqslant a$)，R_3和R_4。对于这个例子中其他的关系容易找到不在关系中的形如(a, a)的对(留给读者作为练习)。■

例9 正整数集合上的“整除”关系是自反的吗？

解 因为只要a是正整数就有$a\mid a$，“整除”关系是自反的。■

在某些关系中一个元素与第二个元素相关，当且仅当第二个元素也与第一个元素相关。比如一个关系由(x, y)对构成，其中x和y是你校的学生并且他们至少学一门公共课程。这个关系就有这种性质。而某些关系有另一种性质，即如果一个元素与第二个元素相关，那么第二个元素就不与第一个元素相关。比如一个关系由(x, y)对构成，其中x和y是你校的学生，且x比y的平均成绩高。这个关系就有后一种性质。

定义4 对于$a, b\in A$，如果只要$(a, b)\in R$就有$(b, a)\in R$，则集合A上的关系R叫做对称的。如果对$a, b\in A$，仅当$a=b$时$(a, b)\in R$和$(b, a)\in R$，则集合A上的关系R叫作反对称的。

注意 使用量词，可以看到如果$\forall a\forall b((a, b)\in R\rightarrow(b, a)\in R)$，则$A$上的关系$R$是对称的。类似地，如果$\forall a\forall b((a, b)\in R\wedge(b, a)\wedge\in R\rightarrow(a=b))$，则$A$上的关系$R$是反对称的。

就是说，关系 R 是对称的，当且仅当如果 a 与 b 相关则 b 与 a 就相关。关系 R 是反对称的，当且仅当不存在由不同元素 a 和 b 构成的有序对使得 a 与 b 相关并且 b 与 a 也相关。对称与反对称的概念不是对立的，因为一个关系可以同时有这两种性质或者两种性质都没有(见节末的练习 8)。一个关系如果包含了某些形如 (a, b) 的对，其中 $a \neq b$，这个关系就不可能同时是对称和反对称的。

注意 尽管从统计数据中可以看出，对称的或反对称的关系相对不多，但许多重要的关系都具有两种性质其中之一(见练习 45)。

例 10 例 7 中的哪些关系是对称的？哪些是反对称的？

解 关系 R_2 和 R_3 是对称的，因为在这些关系中，只要 (a, b) 属于关系就有 (b, a) 属于关系。对于关系 R_2，唯一需要检查的就是(1, 2)和(2, 1)属于这个关系。对于 R_3，必须要检查(1, 2)和(2, 1)属于这个关系，还有(1, 4)和(4, 1)也属于这个关系。读者应该能够验证其他的关系中没有一个是对称的，只需找到一个有序对 (a, b)，使得它在关系中但 (b, a) 不在关系中。

R_4、R_5 和 R_6 都是反对称的。其中每一个关系都不存在这样的有序对，即它由元素 a 和 b 构成，使得 $a \neq b$，但 (a, b) 和 (b, a) 都属于这个关系。读者只需找到有序对 (a, b) 满足 $a \neq b$，但 (a, b) 和 (b, a) 都属于这个关系，就能够验证其他关系中没有一个是反对称的。■

例 11 例 5 中的哪些关系是对称的？哪些是反对称的？

解 关系 R_3、R_4 和 R_6 是对称的。R_3 是对称的，因为如果 $a=b$ 或 $a=-b$，那么就有 $b=a$ 或 $b=-a$。R_4 是对称的，因为 $a=b$ 推出 $b=a$。R_6 是对称的，因为 $a+b \leqslant 3$ 推出 $b+a \leqslant 3$。读者应该能验证其他关系没有一个是对称的。

关系 R_1、R_2、R_4 和 R_5 是反对称的。R_1 是反对称的，因为不等式 $a \leqslant b$ 和 $b \leqslant a$ 推出 $a=b$。R_2 是反对称的，因为 $a>b$ 和 $b>a$ 是不可能的。R_4 是反对称的，因为两个元素相对于 R_4 相关，当且仅当它们是相等的。R_5 是反对称的，因为 $a=b+1$ 和 $b=a+1$ 是不可能的。读者应该能验证其他关系没有一个是反对称的。■

例 12 正整数集合上的整除关系是对称的吗？是反对称的吗？

解 这个关系不是对称的，因为 $1 \mid 2$，但 $2 \nmid 1$。它是反对称的，因为如果 a 和 b 是正整数且 $a \mid b$ 和 $b \mid a$，那么 $a=b$(这个验证留给读者作为练习)。■

设 R 是有序对 (x, y) 构成的关系，其中 x 与 y 是你校的学生，且 x 比 y 得到更多的学分。假设 x 与 y 相关并且 y 与 z 相关，这意味着 x 比 y 得到更多的学分并且 y 比 z 得到更多的学分。可以断言 x 比 z 得到更多的学分，因此 x 与 z 相关。我们证明了 R 有传递性，这个性质定义如下。

定义 5 如果对于 a，b，$c \in A$，$(a, b) \in R$ 并且 $(b, c) \in R$ 则 $(a, c) \in R$，那么集合 A 上的关系 R 叫做传递的。

注意 使用量词，如果 $\forall a \forall b \forall c(((a, b) \in R \wedge (b, c) \in R) \rightarrow (a, c) \in R)$，集合 A 上的关系就叫做传递的。

例 13 例 7 中的关系哪些是传递的？

解 R_4、R_5 和 R_6 是传递的。对于这些关系，我们可以通过验证若 (a, b) 和 (b, c) 属于这个关系则 (a, c) 也属于这个关系来证明每个关系都是传递的。例如，R_4 是传递的，因为只有(3, 2)和(2, 1)，(4, 2)和(2, 1)，(4, 3)和(3, 1)，以及(4, 3)和(3, 2)是这种有序对，而(3, 1)，(4, 1)和(4, 2)属于 R_4。读者应该能验证 R_5 和 R_6 也是传递的。

R_1 不是传递的，因为(3，4)和(4，1)属于 R_1，但(3，1)不属于 R_1。R_2 不是传递的，因为(2，1)和(1，2)属于 R_2，但(2，2)不属于 R_2。R_3 不是传递的，因为(4，1)和(1，2)属于 R_3，但(4，2)不属于 R_3。■

例 14 例 5 中的哪些关系是传递的？

解 关系 R_1、R_2、R_3 和 R_4 是传递的。R_1 是传递的，因为 $a \leqslant b$ 和 $b \leqslant c$ 推出 $a \leqslant c$。R_2 是传递的，因为 $a > b$ 和 $b > c$ 推出 $a > c$。R_3 是传递的，因为 $a = \pm b$ 和 $b = \pm c$ 推出 $a = \pm c$。正如读者应该验证的，显然 R_4 也是传递的。R_5 不是传递的，因为(2，1)和(1，0)属于 R_5，但(2，0)不属于 R_5。R_6 不是传递的，因为(2，1)和(1，2)属于 R_6，但(2，2)不属于 R_6。■

例 15 正整数集合上的"整除"关系是传递的吗？

解 假设 a 整除 b 且 b 整除 c，那么存在正整数 k 和 l 使得 $b = ak$ 和 $c = bl$，因此 $c = a(kl)$，即 a 整除 c。从而证明了这个关系是传递的。■

可以使用计数技术确定具有特殊性质的关系的个数。由此可以得知：这个性质在 n 元素集合的所有关系的集合上有多普遍。

例 16 n 元素集合上有多少个自反的关系？

解 A 上的关系 R 是 $A \times A$ 的子集。因此，要通过指定 $A \times A$ 中 n^2 个有序对中的每一个是否在 R 中来确定关系。因而，如果 R 是自反的，对于 $a \in A$，n 个有序对 (a, a) 中的每一个都必须在 R 中。其他 $n(n-1)$ 个形如 (a, b) 的有序对，$a \neq b$，可能在也可能不在 R 中。因此，由计数的乘积法则，存在 $2^{n(n-1)}$ 个自反的关系。[这就是选择具有 $a \neq b$ 的每个元素 (a, b) 是否属于 R 的方式数。]■

n 元素集合上的对称关系和反对称关系数可以用和例 16 类似的推理得出(见本节末的练习 45)。计数 n 元素集合上的传递关系数的问题已超出本书的范围。

7.1.5 关系的组合

因为从 A 到 B 的关系是 $A \times B$ 的子集，可以按照两个集合组合的任何方式来组合两个从 A 到 B 的关系。考虑例 17 至例 19。

例 17 设 $A = \{1, 2, 3\}$ 和 $B = \{1, 2, 3, 4\}$。组合关系 $R_1 = \{(1, 1), (2, 2), (3, 3)\}$ 和 $R_2 = \{(1, 1), (1, 2), (1, 3), (1, 4)\}$ 可以得到

$$R_1 \cup R_2 = \{(1,1),(1,2),(1,3),(1,4),(2,2),(3,3)\}$$
$$R_1 \cap R_2 = \{(1,1)\}$$
$$R_1 - R_2 = \{(2,2),(3,3)\}$$
$$R_2 - R_1 = \{(1,2),(1,3),(1,4)\}$$
■

例 18 设 A 和 B 分别是学校的所有学生和所有课程的集合。假设 R_1 由所有有序对 (a, b) 组成，其中 a 是选修课程 b 的学生。R_2 由所有的有序对 (a, b) 构成，其中课程 b 是 a 的必修课。关系 $R_1 \cup R_2$，$R_1 \cap R_2$，$R_1 \oplus R_2$，$R_1 - R_2$ 和 $R_2 - R_1$ 是什么？

解 关系 $R_1 \cup R_2$ 由所有的有序对 (a, b) 组成，其中 a 是一个学生，课程 b 要么是他的选修课，要么是他的必修课。$R_1 \cap R_2$ 是有序对 (a, b) 的集合，其中 a 是一个学生，他选修了课程 b 并且课程 b 也是 a 的必修课。$R_1 \oplus R_2$ 由所有的有序对 (a, b) 组成，其中学生 a 已经选修了课程 b 但课程 b 不是 a 的必修课，或者课程 b 是 a 的必修课，但是 a 没有选修它。$R_1 - R_2$ 是所有对 (a, b) 的集合，其中 a 已经选修了课程 b，但 b 不是 a 的必修课，即 b 是 a 的选修课。$R_2 - R_1$ 是所有对 (a, b) 的集合，其中 b 是 a 的必修课，但 a 没有选修它。■

例 19 设 R_1 是实数集合上的"小于"关系，R_2 是实数集合上的"大于"关系，即 $R_1 =$

$\{(x, y) \mid x<y\}$和$R_2=\{(x, y) \mid x>y\}$。$R_1 \cup R_2$，$R_1 \cap R_2$，R_1-R_2，R_2-R_1，$R_1 \oplus R_2$是什么?

解 注意到$(x, y) \in R_1 \cup R_2$当且仅当$(x, y) \in R_1$或$(x, y) \in R_2$，因此$(x, y) \in R_1 \cup R_2$当且仅当$x<y$或$x>y$。由于条件$x<y$或$x>y$与条件$x \neq y$一样，因此$R_1 \cup R_2=\{(x, y) \mid x \neq y\}$。换句话说，"小于"关系与"大于"关系的并是"不相等"关系。

其次，注意到一个对(x, y)同时属于R_1和R_2是不可能的，因为$x<y$并且$x>y$是不可能的。从而得到$R_1 \cap R_2=\varnothing$。也可以看到$R_1-R_2=R_1$，$R_2-R_1=R_2$，$R_1 \oplus R_2=(R_1 \cup R_2)-(R_1 \cap R_2)=\{(x, y) \mid x \neq y\}$。⊖ ■

组合关系有另一种方式，这种方式和函数的复合相似。

定义6 设R是从集合A到集合B的关系，S是从集合B到集合C的关系。R和S的合成是由有序对(a, c)构成的关系，其中$a \in A$，$c \in C$，并且对于它们存在一个元素$b \in B$使得$(a, b) \in R$和$(b, c) \in S$。我们用$S \circ R$表示R与S的合成。

计算两个关系的合成需要我们找出这些元素，它们既是第一个关系中的有序对的第二个元素，也是第二个关系中的有序对的第一个元素。正如例20和例21所示。

例20 R是从$\{1, 2, 3\}$到$\{1, 2, 3, 4\}$的关系且$R=\{(1, 1), (1, 4), (2, 3), (3, 1), (3, 4)\}$，$S$是从$\{1, 2, 3, 4\}$到$\{0, 1, 2\}$的关系且$S=\{(1, 0), (2, 0), (3, 1), (3, 2), (4, 1)\}$，$R$与$S$的合成是什么?

解 $S \circ R$是由所有R中的有序对和S中的有序对构成的，其中在R中有序对的第二元素与S中有序对的第一元素相同。例如，R中的有序对$(2, 3)$和S中的有序对$(3, 1)$产生了$S \circ R$中的有序对$(2, 1)$。计算所有在$R \circ S$中的有序对，我们得到

$$S \circ R=\{(1,0),(1,1),(2,1),(2,2),(3,0),(3,1)\}$$ ■

例21 双亲关系与自身的合成。 设R是所有人集合上的双亲关系，即如果a是b的父母，则$(a, b) \in R$。$(a, c) \in R \circ R$，当且仅当存在一个人b使得a是b的父母且b是c的父母。换句话说，$(a, c) \in R \circ R$当且仅当a是c的祖父母或外祖父母。 ■

从两个关系合成的定义可以递归地来定义关系R的幂。

定义7 设R是集合A上的关系。幂R^n，$n=1, 2, 3, \cdots$，递归地定义为

$$R^1=R \quad 和 \quad R^{n+1}=R^n \circ R$$

这个定义证明了$R^2=R \circ R$，$R^3=R^2 \circ R=(R \circ R) \circ R$，等等。

例22 设$R=\{(1, 1), (2, 1), (3, 2), (4, 3)\}$，求幂$R^n$，$n=2, 3, 4, \cdots$。

解 因为$R^2=R \circ R$，我们得到$R^2=\{(1, 1), (2, 1), (3, 1), (4, 2)\}$。进一步，因为$R^3=R^2 \circ R$，$R^3=\{(1, 1), (2, 1), (3, 1), (4, 1)\}$。其他的计算证明了$R^4$和$R^3$一样，因此$R^4=\{(1, 1), (2, 1), (3, 1), (4, 1)\}$。从而对$n=5, 6, 7, \cdots$，有$R^n=R^3$。读者应该能够验证这个结果。 ■

下面的定理证明一个传递关系的幂是该关系的子集。7.4节将要用到这一结果。

定理1 集合A上的关系R是传递的，当且仅当对$n=1, 2, 3$，有$R^n \subseteq R$。

证 首先证明定理的充分条件。假设对$n=1, 2, 3$，有$R^n \subseteq R$。特别地，有$R^2 \subseteq R$。为证明R的传递性，注意到如果$(a, b) \in R$并且$(b, c) \in R$，根据合成定义就有$(a, c) \in R^2$。因为$R^2 \subseteq R$，这就意味着$(a, c) \in R$。因此R是传递的。

我们将使用数学归纳法证明定理的必要条件。对于$n=1$，定理的这个结果是显而易见的。

⊖ 原书此处为$R_1 \cup R_2-R_1 \cap R_2$，译文根据本书第1章的定义加了括号。——译者注

假定$R^n\subseteq R$，其中n是一个正整数，这是归纳假设。为完成归纳步骤，必须证明这将推出R^{n+1}也是R的子集。为证明这一点，假设$(a, b)\in R^{n+1}$，那么因为$R^{n+1}=R^n\circ R$，存在元素$x\in A$使得$(a, x)\in R$并且$(x, b)\in R^n$。由归纳假设，即$R^n\subseteq R$，推出$(x, b)\in R$。下一步，因为R是传递的，以及$(a, x)\in R$和$(x, b)\in R$，得到$(a, b)\in R$。这就证明了$R^{n+1}\subseteq R$，从而完成了证明。 □

练习

1. 列出从$A=\{0, 1, 2, 3, 4\}$到$B=\{0, 1, 2, 3\}$的关系R中的有序对，其中$(a, b)\in R$，当且仅当
   a) $a=b$　　b) $a+b=4$　　c) $a>b$
   d) $a \mid b$　　e) $\gcd(a, b)=1$　　f) $\text{lcm}(a, b)=2$
2. a) 列出集合$\{1, 2, 3, 4, 5, 6\}$上的关系$R=\{(a, b) \mid a$整除$b\}$中所有的有序对。
   b) 仿照例4用图表示这个关系。
   c) 仿照例4用表表示这个关系。
3. 对集合$\{1, 2, 3, 4\}$上的如下每一个关系，确定它是否是自反的，是否是对称的，是否是反对称的，是否是传递的。
   a) $\{(2, 2), (2, 3), (2, 4), (3, 2), (3, 3), (3, 4)\}$
   b) $\{(1, 1), (1, 2), (2, 1), (2, 2), (3, 3), (4, 4)\}$
   c) $\{(2, 4), (4, 2)\}$
   d) $\{(1, 2), (2, 3), (3, 4)\}$
   e) $\{(1, 1), (2, 2), (3, 3), (4, 4)\}$
   f) $\{(1, 3), (1, 4), (2, 3), (2, 4), (3, 1), (3, 4)\}$
4. 确定所有人的集合上的关系R是否是自反的、对称的、反对称的和传递的，其中$(a, b)\in R$，当且仅当
   a) a比b高。　　b) a和b生在同一天。　　c) a和b同名。　　d) a和b有共同的祖父母。
5. 确定所有Web页上的关系R是否为自反的、对称的、反对称的和传递的，其中$(a, b)\in R$，当且仅当
   a) 每个访问Web页a的人也访问了Web页b。
   b) 在Web页a和b上没有公共链接。
   c) 在Web页a和b上至少有一条公共链接。
   d) 存在一个Web页包含了到Web页a和Web页b的链接。
6. 确定所有实数集合上的关系R是否是自反的、对称的、反对称的和传递的，其中$(x, y)\in R$，当且仅当
   a) $x+y=0$　　b) $x=\pm y$　　c) $x-y$是有理数　　d) $x=2y$
   e) $xy\geqslant 0$　　f) $xy=0$　　g) $x=1$　　h) $x=1$或$y=1$
7. 确定所有整数集合上的关系R是否是自反的、对称的、反对称的和传递的，其中$(x, y)\in R$，当且仅当
   a) $x\neq y$　　b) $xy\geqslant 1$　　c) $x=y+1$或$x=y-1$　　d) $x\equiv y \pmod 7$
   e) x是y的倍数　　f) x与y都是负的或都是非负的　　g) $x=y^2$　　h) $x\geqslant y^2$
8. 给出一个集合上的关系的例子，要求它是
   a) 对称的和反对称的。　　b) 既不是对称的也不是反对称的。

☞ 如果对于每个$a\in A$，有$(a, a)\notin R$，那么集合A上的关系R是**反自反的**，即如果没有A中的元素与自己有关系，关系R就是反自反的。

9. 练习3的哪些关系是反自反的？
10. 练习4的哪些关系是反自反的？
11. 练习5的哪些关系是反自反的？
12. 练习6的哪些关系是反自反的？
13. 集合上的关系可能既不是自反的也不是反自反的吗？

14. 使用量词表达一个关系是反自反的含义。
15. 给出在所有人的集合上一个反自反关系的例子。

一个关系 R 叫做非对称的，如果 $(a, b)\in R$ 推出 $(b, a)\notin R$。

16. 练习 3 的哪些关系是非对称的？
17. 练习 4 的哪些关系是非对称的？
18. 练习 5 的哪些关系是非对称的？
19. 练习 6 的哪些关系是非对称的？
20. 非对称的关系一定是反对称的吗？反对称的关系一定是非对称的吗？对你的答案说明理由。
21. 使用量词表达一个关系是非对称的含义。
22. 给出在所有人的集合上一个非对称关系的例子。
23. 从 m 元素集合到 n 元素集合有多少个不同的关系？

☞ 设 R 是从集合 A 到集合 B 的关系。从 B 到 A 的逆关系是有序对的集合 $\{(b, a) \mid (a, b)\in R\}$，记作 R^{-1}，补关系 $\bar{R}$ 是有序对的集合 $\{(a, b) \mid (a, b)\notin R\}$。

24. 设 R 是整数集合上的关系 $R=\{(a, b) \mid a<b\}$，求
    a) R^{-1}　　b) $\bar{R}$
25. 设 R 是正整数集合上的关系 $R=\{(a, b) \mid a$ 整除 $b\}$，求
    a) R^{-1}　　b) $\bar{R}$
26. 设 R 是美国所有州的集合上的关系，R 由有序对 (a, b) 构成，其中 a 州与 b 州相邻接，求
    a) R^{-1}　　b) $\bar{R}$
27. 设从 A 到 B 的函数 f 是一一对应的。令 R 是和 f 的图相等的关系，即 $R=\{(a, f(a)) \mid a\in A\}$。逆关系 R^{-1} 是什么？
28. 令 $R_1=\{(1, 2), (2, 3), (3, 4)\}$，$R_2=\{(1, 1), (1, 2), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3), (3, 4)\}$ 是从 $\{1, 2, 3\}$ 到 $\{1, 2, 3, 4\}$ 的关系，求
    a) $R_1\cup R_2$　　b) $R_1\cap R_2$　　c) R_1-R_2　　d) R_2-R_1
29. 设 A 是你校的学生的集合，B 是学校图书馆的书的集合。设 R_1 和 R_2 都是有序对 (a, b) 构成的关系，在 R_1 中，学生 a 修一门课程需要读书 b，在 R_2 中，学生 a 已经读过书 b。描述在下面每个关系中的有序对。
    a) $R_1\cup R_2$　　b) $R_1\cap R_2$　　c) $R_1\oplus R_2$　　d) R_1-R_2　　e) R_2-R_1
30. 设 R 是关系 $\{(1, 2), (1, 3), (2, 3), (2, 4), (3, 1)\}$，$S$ 是关系 $\{(2, 1), (3, 1), (3, 2), (4, 2)\}$，求 $S\circ R$。
31. 设 R 是由人的集合上的有序对 (a, b) 组成的关系，其中 a 是 b 的父母。设 S 是由人的集合上的有序对 (a, b) 组成的关系，其中 a 是 b 的兄弟姐妹。$S\circ R$ 和 $R\circ S$ 是什么关系？

练习 32～35 涉及的都是实数集合上的关系：

$R_1=\{(a, b)\in \mathbf{R}^2 \mid a>b\}$，“大于”关系

$R_2=\{(a, b)\in \mathbf{R}^2 \mid a\geqslant b\}$，“大于或等于”关系

$R_3=\{(a, b)\in \mathbf{R}^2 \mid a<b\}$，“小于”关系

$R_4=\{(a, b)\in \mathbf{R}^2 \mid a\leqslant b\}$，“小于或等于”关系

$R_5=\{(a, b)\in \mathbf{R}^2 \mid a=b\}$，“等于”关系

$R_6=\{(a, b)\in \mathbf{R}^2 \mid a\neq b\}$，“不等”关系

32. 求
    a) $R_1\cup R_3$　　b) $R_1\cup R_5$　　c) $R_2\cap R_4$　　d) $R_3\cap R_5$
    e) R_1-R_2　　f) R_2-R_1　　g) $R_1\oplus R_3$　　h) $R_2\oplus R_4$
33. 求

a) $R_2 \cup R_4$　b) $R_3 \cup R_6$　c) $R_3 \cap R_6$　d) $R_4 \cap R_6$
e) $R_3 - R_6$　f) $R_6 - R_3$　g) $R_2 \oplus R_6$　h) $R_3 \oplus R_5$

34. 求
a) $R_1 \circ R_1$　b) $R_1 \circ R_2$　c) $R_1 \circ R_3$　d) $R_1 \circ R_4$
e) $R_1 \circ R_5$　f) $R_1 \circ R_6$　g) $R_2 \circ R_3$　h) $R_3 \circ R_3$

35. 求
a) $R_2 \circ R_1$　b) $R_2 \circ R_2$　c) $R_3 \circ R_5$　d) $R_4 \circ R_1$
e) $R_5 \circ R_3$　f) $R_3 \circ R_6$　g) $R_4 \circ R_6$　h) $R_6 \circ R_6$

36. 设 R 是所有人的集合上的双亲关系(见例 21)。一个有序对在关系 R^3 中?

37. 设 R 是具有博士学位的人的集合上的关系,要求$(a, b) \in R$,当且仅当 a 是 b 的导师。什么情况下一个有序对(a, b)在 R^2 中?什么情况下一个有序对(a, b)在 R^n 中?这里 n 是正整数。(注意每个具有博士学位的人有一个导师。)

38. 设 R_1 和 R_2 分别是所有正整数集合上的“整除”和“倍数”关系,即 $R_1 = \{(a, b) \mid a \text{ 整除 } b\}$和 $R_2 = \{(a, b) \mid a \text{ 是 } b \text{ 的倍数}\}$。求
a) $R_1 \cup R_2$　b) $R_1 \cap R_2$　c) $R_1 - R_2$　d) $R_2 - R_1$　e) $R_1 \oplus R_2$

39. 设 R_1 和 R_2 分别是整数集合上的“模 3 同余”和“模 4 同余”关系,即 $R_1 = \{(a, b) \mid a \equiv b \pmod 3\}$和$R_2 = \{(a, b) \mid a \equiv b \pmod 4\}$。求
a) $R_1 \cup R_2$　b) $R_1 \cap R_2$　c) $R_1 - R_2$　d) $R_2 - R_1$　e) $R_1 \oplus R_2$

40. 列出集合$\{0, 1\}$上的 16 个不同的关系。

41. $\{0, 1\}$上的 16 个不同的关系中有多少个包含了对$(0, 1)$?

42. 你在练习 40 列出的$\{0, 1\}$集合上的 16 个关系中哪些是
a) 自反的?　b) 反自反的?　c) 对称的?　d) 反对称的?
e) 非对称的?　f) 传递的?

43. a) 在集合$\{a, b, c, d\}$上有多少个关系?
b) 在集合$\{a, b, c, d\}$上有多少个关系包含对(a, a)?

44. 设 S 是 n 个元素的集合,a 和 b 是 S 中的不同元素。S 上有多少个关系使得
a) (a, b)属于这个关系?
b) (a, b)不属于这个关系?
c) 在关系中没有有序对以 a 作为它们的第一元素?
d) 在关系中至少有一个有序对以 a 作为第一元素?
e) 在关系中没有有序对以 a 作为它们的第一元素,也没有有序对以 b 作为它们的第二元素?
f) 在关系中至少有一个有序对以 a 作为它的第一元素或者以 b 作为它的第二元素?

*45. n 元素集合上有多少个关系是
a) 对称的?　b) 反对称的?　c) 非对称的?　d) 反自反的?　e) 自反的和对称的?
f) 既不是自反的也不是反自反的?

*46. n 元素集合上有多少个传递的关系?如果
a) $n=1$　b) $n=2$　c) $n=3$

47. 找出在下面定理证明中的错误。
“定理”:设 R 是集合 A 上的对称和传递的关系,则 R 是自反的。
“证明”:设 $a \in A$。取元素 $b \in A$ 使得$(a, b) \in R$。由于 R 是对称的,因而有$(b, a) \in R$。现在使用传递性,由$(a, b) \in R$ 和$(b, a) \in R$ 可以得出结论$(a, a) \in R$。

48. 假设 R 和 S 是集合 A 上的自反关系。证明或反驳下面的每个论断。
a) $R \cup S$ 是自反的　b) $R \cap S$ 是自反的　c) $R \oplus S$ 是反自反的
d) $R - S$ 是反自反的　e) $S \circ R$ 是自反的

49. 证明集合 A 上的关系 R 是对称的，当且仅当 $R=R^{-1}$，其中 R^{-1} 是 R 的逆关系。
50. 证明集合 A 上的关系 R 是反对称的，当且仅当 $R\cap R^{-1}$ 是对角线关系 $\Delta=\{(a, a) \mid a\in A\}$ 的子集。
51. 证明集合 A 上的关系 R 是自反的，当且仅当逆关系 R^{-1} 是自反的。
52. 证明集合 A 上的关系 R 是自反的，当且仅当补关系 $\bar{R}$ 是反自反的。
53. 设 R 是自反的和传递的关系，证明对所有的正整数 n，$R^n=R$。
54. 设 R 是集合{1，2，3，4，5}上的关系，R 包含有序对(1，1)，(1，2)，(1，3)，(2，3)，(2，4)，(3，1)，(3，4)，(3，5)，(4，2)，(4，5)，(5，1)，(5，2)和(5，4)，求
   a) R^2 b) R^3 c) R^4 d) R^5
55. 设 R 是集合 A 上的自反关系，证明对所有的正整数 n，R^n 也是自反的。
*56. 设 R 是对称关系，证明对所有的正整数 n，R^n 也是对称的。
57. 假设关系 R 是反自反的，R^2 一定是反自反的吗？对你的答案给出理由。

7.2 n 元关系及其应用

7.2.1 引言

在两个以上集合的元素中常常会产生某种关系。例如，学生的姓名、学生的专业以及学生的平均学分绩点之间的关系。类似地，一个航班的航空公司、航班号、出发地、目的地、起飞时间和到达时间等也有一种关系。在数学中也有这种关系。例如有 3 个整数，其中第一个整数比第二个整数大，而第二个整数比第三个整数大。另一个例子是直线上的点之间的关系，即当第二个点在第一和第三个点之间时，这三个点有关系。

本节我们将研究两个以上集合的元素之间的关系。这种关系叫 *n 元关系*。可以用这种关系表示计算机的数据库。这种表示有助于回答对数据库中所存信息的查询，例如：哪个航班在午夜 3 点到 4 点之间降落在 O'Hare 机场？你们学校的二年级学生哪些是主修数学或计算机科学的，并且平均学分绩点大于 3.0？公司的哪些雇员为这个公司工作不到 5 年但所赚的钱已经超过 50 000 美元？

7.2.2 n 元关系

我们从建立关系数据库理论的基本定义开始。

定义 1 设 $A_1, A_2, \cdots, A_n$ 是集合。在这些集合上的 n 元关系是 $A_1\times A_2\times\cdots\times A_n$ 的子集。这些集合 $A_1, A_2, \cdots, A_n$ 叫做关系的域，n 叫做它的阶。

例 1 R 是由三元组(a, b, c)构成的关系，其中 a、b、c 是满足 $a<b<c$ 的整数。那么 $(1, 2, 3)\in R$，但$(2, 4, 3)\notin R$。这个关系的阶是 3。它的域都等于整数集合。 ■

例 2 设 R 是由 5 元组(A, N, S, D, T)构成的表示飞机航班的关系，其中 A 是航空公司，N 是航班号，S 是出发地，D 是目的地，T 是起飞时间。例如，如果 Nadir 直达航空公司在 15:00 有从 Newark 到 Bangor 的 963 航班，那么(Nadir，963，Newark，Bangor，15:00)属于 R。这个关系的阶是 5，并且它的域是所有航空公司的集合、航班号的集合、城市的集合、城市的集合以及时间的集合。 ■

7.2.3 数据库和关系



操作数据库信息所需要的时间依赖于这些信息是怎样存储的。插入和删除记录，更新记录，检索记录以及从一个重叠的数据库中组合记录的操作，在一个大型数据库中每天要执行几百万次。由于这些操作的重要性，已经开发了数据库表示的各种方法。我们将讨

论其中的一种基于关系概念的方法，叫做**关系数据模型**。

数据库由记录组成，这些记录是由字段构成的 n 元组。这些字段是 n 元组的数据项。例如，学生记录的数据库可以由包含学生的姓名、学号、专业、平均学分绩点(GPA)的字段构成。关系数据模型把一个记录的数据库表示成一个 n 元关系。于是，学生记录可以被表示成形如(学生姓名，学号，专业，GPA)的 4 元组。6 个记录的一个数据库样本是：

(Ackermann，231455，计算机科学，3.88)

(Adams，888323，物理学，3.45)

(Chou，102147，计算机科学，3.49)

(Goodfriend，453876，数学，3.45)

(Rao，678543，数学，3.90)

(Stevens，786576，哲学，2.99)

用于表示数据库的关系叫做**表**，因为这些关系常常用表来给出。这些表的每个列对应于数据库的一个**属性**。例如，同样的学生数据库显示在表 7-1。这个数据库的属性是学生姓名、学号、专业和平均学分绩点(GPA)。

表 7-1　学生

学生姓名	学号	专业	GPA
Ackermann	231 455	计算机科学	3.88
Adams	888 323	物理学	3.45
Chou	102 147	计算机科学	3.49
Goodfriend	453 876	数学	3.45
Rao	678 543	数学	3.90
Stevens	786 576	心理学	2.99

当 n 元组的某个域的值能够确定这个 n 元组时，n 元关系的这个域就叫做**主关键字**。这就是说当关系中没有两个 n 元组在这个域有相同的值时，这个域就是主关键字。

常常要从数据库增加或删除记录。由于这一点，一个域是主关键字的性质是随时间而改变的。所以，一个主关键字应该选择那种无论数据库怎样改变都能继续存在的字段。用数据库的**内涵**的主关键字就可以做到这一点，它包含了在表示这个数据库的 n 元关系中可能永远含有的所有的 n 元组。

例 3　假设将来不再增加 n 元组，对于表 7-1 所示的 n 元关系，哪个域是主关键字？

解　因为在这个表中对应每个学生的姓名只有一个 4 元组，学生姓名的域是主关键字。类似地，在这个表中学号是唯一的，学号的域也是主关键字。但是，所学专业的域不是主关键字，因为包含同样的专业的 4 元组有多个。平均学分绩点的域也不是主关键字，因为有 2 个 4 元组包含了同样的 GPA。 ■

在一个 n 元关系中域的组合也可以唯一地标识 n 元组。当一组域的值确定了一个关系中的 n 元组时，这些域的笛卡儿积就叫做**复合关键字**。

例 4　对于表 7-1 中的 n 元关系，假设不增加 n 元组，所学专业的域与 GPA 的域的笛卡儿积是复合关键字吗？

解　这个表中没有两个 4 元组有同样的专业和同样的 GPA，因此这个笛卡儿积是一个复合关键字。 ■

因为主关键字和复合关键字用于唯一地标识数据库中的记录，当新的记录被加到这个数据库时关键字要保持有效是非常重要的。因此，应该做检测以保证在这个或这些相应的字段中每个新记录与表中所有其他的记录不同。例如，使用学号作为学生记录的关键字是有意义的，因为没有两个学生有同样的学号。一个大学不应该使用姓名字段作为关键字，因为有可能两个学生有同样的姓名(如 John Smith)。

7.2.4 n 元关系的运算

n 元关系存在可以用于构成新的 n 元关系的不同的运算。综合使用这些运算能够回答对数据库的满足特定条件的所有元组的查询。

一个 n 元关系最基本的运算是在这个 n 元关系中确定满足特定条件的所有 n 元组。例如，我们可能想在学生记录的数据库中找出主修计算机科学专业的所有学生的记录。可能我们想在同一个数据库中找出所有平均学分绩点在 3.5 以上的学生。也可能我们想在这个数据库中找出所有主修计算机科学专业的平均学分绩点在 3.5 以上的学生。为完成这些任务我们使用了选择运算。

定义 2　设 R 是 n 元关系，C 是 R 中元素可能满足的一个条件。那么选择运算 s_C 将 n 元关系 R 映射到 R 中满足条件 C 的所有 n 元组构成的 n 元关系。

例 5　为了找出表 7-1 所示的 n 元关系 R 中主修计算机科学专业的学生的记录，我们使用运算 s_{C_1}，其中 C_1 是条件专业＝"计算机科学"。结果是两个 4 元组(Ackermann，231455，计算机科学，3.88)和(Chou，102147，计算机科学，3.49)。类似地，为了在这个数据库中找出平均学分绩点在 3.5 以上的学生记录，我们使用运算 s_{C_2}，其中 C_2 是条件 GPA>3.5。结果是两个 4 元组(Ackermann，231455，计算机科学，3.88)和(Rao，678543，数学，3.90)。最后，为找出主修计算机科学专业的 GPA 在 3.5 以上的学生记录，我们使用运算 s_{C_3}，其中 C_3 是条件(专业＝"计算机科学"∧GPA>3.5)。结果由一个 4 元组(Ackermann，231455，计算机科学，3.88)构成。■

通过投影删去关系中每条记录相同的域可以构建一个新的 n 元关系。

定义 3　投影 $P_{i_1,i_2,\cdots,i_m}$ 将 n 元组(a_1，a_2，…，a_n)映到 m 元组(a_{i_1}，a_{i_2}，…，a_{i_m})，其中 $m\leqslant n$。

换句话说，投影 $P_{i_1,i_2,\cdots,i_m}$ 删除了 n 元组的 $n-m$ 个分量，保留了第 i_1，i_2，…，i_m 个分量。

例 6　当投影 $P_{1,3}$ 施用到 4 元组(2，3，0，4)，(Jane Doe，234111001，地理学，3.14)以及(a_1，a_2，a_3，a_4)时结果是什么？

解　$P_{1,3}$ 把这些 4 元组分别映到(2，0)，(Jane Doe，地理学)和(a_1，a_3)。■

例 7 说明了怎样使用投影来产生新关系。

例 7　当投影 $P_{1,4}$ 施用于表 7-1 中的关系时结果是什么？

解　当施用投影 $P_{1,4}$ 时，表的第二列和第三列被删除，得到了表示学生姓名和平均学分绩点的对。表 7-2 给出了这个投影的结果。■

当一个投影施用到一个关系的表中时，有可能使行变少。当关系中的某些 n 元组在投影的 m 个分量中每个分量的值都相同，只在被删除的分量有不同的值时，就会出现这种情况。例如，考虑下面的例子。

例 8　当投影 $P_{1,2}$ 施用到表 7-3 的关系时得到什么表？

解　表 7-4 给出了当投影 $P_{1,2}$ 施用到表 7-3 时得到的关系。注意在施用了这个投影后行数减少。■

表 7-2 GPA

学生姓名	GPA
Ackermann	3.88
Adams	3.45
Chou	3.49
Goodfriend	3.45
Rao	3.90
Stevens	2.99

表 7-3 注册

学生	专业	课程
Glauser	生物学	BI290
Glauser	生物学	MS475
Glauser	生物学	PY410
Marcus	数学	MS511
Marcus	数学	MS603
Marcus	数学	CS322
Miller	计算机科学	MS575
Miller	计算机科学	CS455

当两个表分享着某些相同的字段时连接运算将这两个表组合成一个表。例如，一个表包含了航空公司、航班号和通道的字段，另一个表包含了航班号、通道和起飞时间的字段。可以将这两个表组合成一个包含航空公司、航班号、通道和起飞时间字段的表。

表 7-4 专业

学生	专业
Glauser	生物学
Marcus	数学
Miller	计算机科学

定义 4 设 R 是 m 元关系且 S 是 n 元关系，其中 $p\leqslant m$ 和 $p\leqslant n$。连接 $J_p(R, S)$ 是 $m+n-p$ 元关系，它包含了所有的 $(m+n-p)$ 元组 $(a_1, a_2, \cdots, a_{m-p}, c_1, c_2, \cdots, c_p, b_1, b_2, \cdots, b_{n-p})$，其中 m 元组 $(a_1, a_2, \cdots, a_{m-p}, c_1, c_2, \cdots, c_p)$ 属于 R 且 n 元组 $(c_1, c_2, \cdots, c_p, b_1, b_2, \cdots, b_{n-p})$ 属于 S。

换句话说，连接运算 J_p 从两个关系产生一个新的关系，它把第一个关系的所有 m 元组和第二个关系的所有 n 元组组合起来，其中 m 元组的后 p 个分量与 n 元组的前 p 个分量相同。

例 9 当用连接运算 J_2 组合表 7-5 和表 7-6 中的关系时所得结果关系是什么？

解 J_2 产生的关系在表 7-7 中。 ■

表 7-5 教学课程

教授	系	课号
Cruz	动物学	335
Cruz	动物学	412
Farber	心理学	501
Farber	心理学	617
Grammer	物理学	544
Grammer	物理学	551
Rosen	计算机科学	518
Rosen	数学	575

表 7-6 教室安排

系	课号	教室	时间
计算机科学	518	N521	2:00P.M.
数学	575	N502	3:00P.M.
数学	611	N521	4:00P.M.
物理学	544	B505	4:00P.M.
心理学	501	A100	3:00P.M.
心理学	617	A110	11:00A.M.
动物学	335	A100	9:00A.M.
动物学	412	A100	8:00A.M.

表 7-7 教学安排

教授	系	课号	教室	时间
Cruz	动物学	335	A100	9:00A.M.
Cruz	动物学	412	A100	8:00A.M.
Farber	心理学	501	A100	3:00P.M.
Farber	心理学	617	A110	11:00A.M.
Grammer	物理学	544	B505	4:00P.M.
Rosen	计算机科学	518	N521	2:00P.M.
Rosen	数学	575	N502	3:00P.M.

从已知关系产生新关系的运算除了投影和连接运算以外还有其他运算。对这些运算的描述可以在关于数据库理论的书中找到。

7.2.5 SQL

数据库查询语言 SQL(结构化查询语言 Structured Query Language 的缩写)可以用来执行本节所描述的运算。例 10 说明了 SQL 命令是怎样与 n 元关系上的运算相关的。

例 10 我们将说明怎样用 SQL 来表达查询，这是通过显示怎样使用 SQL 从表 7-8 中做一次关于航班的查询来说明的。SQL 语句

```
SELECT Departure_time             (选择起飞时间)
FROM Flights                      (从航班)
WHERE destination = 'Detroit'     (当目的地 = '底特律')
```

是用于在航班数据库中找出投影 P_5，它能够(在起飞时间属性上)选出 5 元组使得这些 5 元组满足条件：destination = 'Detroit'。输出是一张以底特律为目的地的包含航班时间的表，即 08:10，08:47和 09:44。SQL 使用 FROM 子句识别查询施用到的 n 元关系，WHERE 子句说明选择运算的条件，而 SELECT 子句说明将被使用的投影运算。(小心：SQL 使用 SELECT 表示一个投影，而不是一个选择运算。这是一个令人遗憾的术语不一致的例子。) ■

表 7-8　航班

航空公司	航班号	通道	目的地	起飞时间
Nadir	122	34	底特律	08:10
Acme	221	22	丹佛	08:17
Acme	122	33	安克雷奇	08:22
Acme	323	34	檀香山	08:30
Nadir	199	13	底特律	08:47
Acme	222	22	丹佛	09:10
Nadir	322	34	底特律	09:44

例 11 说明 SQL 怎样做涉及多个表的查询。

例 11 SQL 语句

```
SELECT professor, time                              (选择教授，时间)
FROM teaching_assignments, class_schedule           (从教学课程，教室安排)
WHERE department = 'mathematics'                    (当系 = '数学')
```

用于找出在数据库(如表 7-7 所示)中的 5 元组的投影 $P_{1,5}$，这个数据库是分别在表 7-5 中的教学课程和表 7-6 中的教室安排，数据库满足 department = 'mathematics'条件下的连接 J_2。输出由单一的 2 元组(Rosen，3:00PM)构成。这里的 SQLFROM 子句用于找出两个不同数据库的连接。 ■

本节我们仅仅接触到关系数据库的基本概念。更多的信息可以在[AhU195]中找到。

练习

1. 列出关系$\{(a, b, c) \mid a$，b 和 c 是整数且 $0<a<b<c<5\}$中的三元组。
2. 在关系$\{(a, b, c, d) \mid a$，b，c，d 是正整数且 $abcd=6\}$中有哪些 4 元组？
3. 列出表 7-8 关系中的 5 元组。

4. 假设不增加新的 n 元组，对下面表中的关系找出所有的主关键字。

   a) 表 7-3　　b) 表 7-5　　c) 表 7-6　　d) 表 7-8

5. 假设不增加新的 n 元组，对于表 7-8 中的数据库找出一个由两个字段构成的复合关键字，其中一个字段是航空公司。
6. 假设不增加新的 n 元组，对于表 7-7 中的数据库找出一个由两个字段构成的复合关键字，其中一个字段是教授。
7. 3 元关系中的 3 元组表示了一个学生数据库中的下述属性：学号、姓名、电话号码。
   a) 学号可能是主关键字吗？
   b) 姓名可能是主关键字吗？
   c) 电话号码可能是主关键字吗？
8. 4 元关系中的 4 元组表示了出版图书的下述属性：书名、书号、出版日期、页数。
   a) 什么可能是这个关系的主关键字？
   b) 在什么条件下(书名、出版日期)是复合关键字？
   c) 在什么条件下(书名、页数)是复合关键字？
9. 5 元关系中的 5 元组表示了美国所有人的下述属性：姓名、社会保险号、住址、城市、州。
   a) 对这个关系确定一个主关键字。
   b) 在什么条件下(姓名、住址)是复合关键字？
   c) 在什么条件下(姓名、住址、城市)是复合关键字？
10. 设 C 是条件：教室＝A100，当你施用选择运算 s_C 到表 7-7 的数据库时，可以得到什么？
11. 设 C 是条件：目的地＝底特律，当你施用选择运算 s_C 到表 7-8 的数据库时，可以得到什么？
12. 设 C 是条件：(项目＝2) ∧ (数量＞50)，当你施用选择运算 s_C 到表 7-10 的数据库时，可以得到什么？
13. 设 C 是条件：(航空公司＝Nadir) ∨ (目的地＝丹佛)，当你施用选择运算 s_C 到表 7-8 的数据库时，可以得到什么？
14. 当施用投影 $P_{2,3,5}$ 到 5 元组(a，b，c，d，e)时，你能得到什么？
15. 哪个投影映射用于删除一个 6 元组的第一、第二和第四个分量？
16. 给出施用投影 $P_{1,2,4}$ 到表 7-8 以后得到的表。
17. 给出施用投影 $P_{1,4}$ 到表 7-8 以后得到的表。
18. 把连接运算 J_3 施用到 5 元组的表和 8 元组的表后所得表中的 n 元组里有多少个分量？
19. 构造把连接运算 J_2 施用到表 7-9 和表 7-10 的关系中所得到的表。

表 7-9　零件需求

供货商	零件号	项目
23	1092	1
23	1101	3
23	9048	4
31	4975	3
31	3477	2
32	6984	4
32	9191	2
33	1001	1

表 7-10　零件库存

零件号	项目	数量	颜色代码
1001	1	14	8
1092	1	2	2
1101	3	1	1
3477	2	25	2
4975	3	6	2
6984	4	10	1
9048	4	12	2
9191	2	80	4

20. 证明如果 C_1 和 C_2 是 n 元关系 R 的元素可能满足的条件，那么 $s_{C_1 \wedge C_2}(R)=s_{C_1}(s_{C_2}(R))$。
21. 证明如果 C_1 和 C_2 是 n 元关系 R 的元素可能满足的条件，那么 $s_{C_1}(s_{C_2}(R))=s_{C_2}(s_{C_1}(R))$。
22. 证明如果 C 是 n 元关系 R 和 S 的元素可能满足的条件，那么 $s_C(R \cup S)=s_C(R) \cup s_C(S)$。
23. 证明如果 C 是 n 元关系 R 和 S 的元素可能满足的条件，那么 $s_C(R \cap S)=s_C(R) \cap s_C(S)$。

24. 证明如果 C 是 n 元关系 R 和 S 的元素可能满足的条件，那么 $s_C(R-S)=s_C(R)-s_C(S)$。
25. 证明如果 R 和 S 是 n 元关系，那么 $P_{i_1,i_2,\cdots,i_m}(R\cup S)=P_{i_1,i_2,\cdots,i_m}(R)\cup P_{i_1,i_2,\cdots,i_m}(S)$。
26. 给出一个例子证明如果 R 和 S 是两个 n 元关系，那么 $P_{i_1,i_2,\cdots,i_m}(R\cap S)$ 可能与 $P_{i_1,i_2,\cdots,i_m}(R)\cap P_{i_1,i_2,\cdots,i_m}(S)$ 不同。
27. 给出一个例子证明如果 R 和 S 是两个 n 元关系，那么 $P_{i_1,i_2,\cdots,i_m}(R-S)$ 可能与 $P_{i_1,i_2,\cdots,i_m}(R)-P_{i_1,i_2,\cdots,i_m}(S)$ 不同。
28. a) 与下述用 SQL 语句表示的查询相对应的运算是什么？

```
SELECT Supplier                    (选择供货商)
FROM Part _ needs                  (从零件需求)
WHERE 1000≤Part _ number≤5000      (当 1000≤零件号≤5000)
```

b) 给定表 7-9 的数据库作为输入，这个查询的输出是什么？

29. a) 与下述用 SQL 语句表示的查询相对应的运算是什么？

```
SELECT Supplier, Project                 (选择供货商，项目)
FROM Part _ needs, part _ inventory      (从零件需求，零件库存)
WHERE Quantity≤10                        (当数量≤10)
```

b) 给定表 7-9 和表 7-10 的数据库作为输入，这个查询的输出是什么？

7.3 关系的表示

7.3.1 引言

有多种方式表示有穷集之间的关系。正如我们已经看到的，一种方式就是列出它的有序对。本节将讨论另外两种表示关系的方式。一种方式就是使用 0-1 矩阵。另一种方式就是使用有向图。

一般说来，矩阵适用于计算机程序中关系的表示。另一方面，人们常常发现使用有向图来表示关系对理解这些关系的性质是很有用的。

7.3.2 用矩阵表示关系

可以用 0-1 矩阵来表示一个有穷集之间的关系。假设 R 是从 $A=\{a_1, a_2, \cdots, a_m\}$ 到 $B=\{b_1, b_2, \cdots, b_n\}$ 的关系。（这里集合 A 和集合 B 的元素已经按照某一特定的但是任意的次序列出。此外，当 $A=B$ 时我们对于 A 和 B 使用同样的排序。）关系 R 可以用矩阵 $\mathbf{M}_R=[m_{ij}]$ 来表示，其中

$$m_{ij}=\begin{cases}1 & \text{如果}(a_i,b_j)\in R\\ 0 & \text{如果}(a_i,b_j)\notin R\end{cases}$$

换句话说，当 a_i 和 b_j 有关系时表示 R 的 0-1 矩阵的 (i, j) 项是 1，当 a_i 和 b_j 没关系时在这个位置是 0（这种表示依赖于对 A 和 B 使用的排序）。

下面的例子说明用矩阵来表示关系。

例 1 假设 $A=\{1, 2, 3\}$，$B=\{1, 2\}$。令 R 是从 A 到 B 的关系，且如果 $a\in A$，$b\in B$，$a>b$，则 R 包含 (a, b)。如果 $a_1=1$，$a_2=2$，$a_3=3$，$b_1=1$，$b_2=2$，表示 R 的矩阵是什么？

解 因为 $R=\{(2, 1), (3, 1), (3, 2)\}$，关于 R 的矩阵是

$$\mathbf{M}_R=\begin{bmatrix}0 & 0\\ 1 & 0\\ 1 & 1\end{bmatrix}$$

$\mathbf{M}_R$ 中的 1 说明了有序对(2，1)，(3，1)和(3，2)属于 R，0 说明了没有其他的有序对属于 R。 ■

例 2 设 $A=\{a_1, a_2, a_3\}$，$B=\{b_1, b_2, b_3, b_4, b_5\}$。哪些有序对在下面的矩阵所表示的关系 R 中？

$$\mathbf{M}_R = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

解 因为 R 由使得 $m_{ij}=1$ 的有序对(a_i, b_j)构成，因而

$$R = \{(a_1, b_2), (a_2, b_1), (a_2, b_3), (a_2, b_4), (a_3, b_1), (a_3, b_3), (a_3, b_5)\}$$ ■

集合上的关系的矩阵是一个方阵，可以用这个矩阵确定关系是否有某种性质。A 上的关系是自反的，如果对于每个 $a \in A$ 有$(a, a) \in R$。于是，R 是自反的，当且仅当对 $i=1, 2, \cdots, n$，$(a_i, a_i) \in R$。也就有，R 是自反的，当且仅当对 $i=1, 2, \cdots, n$，$m_{ii}=1$。换句话说，如果如图 7-3 所示，$\mathbf{M}_R$ 的主对角线的所有元素都等于 1，那么 R 是自反的。

如果$(a, b) \in R$ 推出$(b, a) \in R$，那么关系 R 是对称的。因此集合 $A=\{a_1, a_2, \cdots, a_n\}$上的关系 R 是对称的，当且仅当只要$(a_i, a_j) \in R$ 就有$(a_j, a_i) \in R$。用矩阵 $\mathbf{M}_R$ 的术语来说，R 是对称的当且仅当只要 $m_{ij}=1$ 就有 $m_{ji}=1$。这也意味着只要 $m_{ij}=0$ 就有 $m_{ji}=0$。因此 R 是对称的，当且仅当对所有的整数对 i，j(其中 $i=1, 2, \cdots, n$，$j=1, 2, \cdots, n$)都有 $m_{ij}=m_{ji}$。回顾 2.7 节矩阵转置的定义，我们看到 R 是对称的当且仅当

$$\mathbf{M}_R = (\mathbf{M}_R)^t$$

即 $\mathbf{M}_R$ 是对称矩阵。对称关系的矩阵形式如图 7-4a 所示。

图 7-3 关于自反关系的 0-1 矩阵

a) 对称的　　b) 反对称的

图 7-4 关于对称和反对称关系的 0-1 矩阵

关系 R 是反对称的，当且仅当$(a, b) \in R$ 和$(b, a) \in R$ 推出 $a=b$。因此，反对称关系的矩阵有下述性质，即如果 $m_{ij}=1$，$i \neq j$，则 $m_{ji}=0$。或者换句话说，当 $i \neq j$ 时，$m_{ij}=0$ 或 $m_{ji}=0$。反对称关系的矩阵形式显示在图 7-4b 中。

例 3 假设集合上的关系 R 由矩阵

$$\mathbf{M}_R = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

表示，R 是自反的、对称的和反对称的吗？

解 因为这个矩阵所有的对角线元素都等于 1，R 是自反的。又由于 $\mathbf{M}_R$ 是对称的，因此 R 是对称的。也容易看出 R 不是反对称的。 ■

可以用布尔运算加和乘(在 2.7 节讨论)找到两个关系的并和交的矩阵表示。假设集合 A 上的关系 R_1 和 R_2 分别由矩阵 $\mathbf{M}_{R_1}$ 和 $\mathbf{M}_{R_2}$ 来表示。$\mathbf{M}_{R_1}$ 或 $\mathbf{M}_{R_2}$ 在某个位置为 1，则关系的并的矩阵相应位置有 1。$\mathbf{M}_{R_1}$ 和 $\mathbf{M}_{R_2}$ 在某个位置同时为 1，则关系的交的矩阵相应位置有 1。于是，这些关系的并和交的矩阵表示是

$$\mathbf{M}_{R_1\cup R_2}=\mathbf{M}_{R_1}\vee\mathbf{M}_{R_2}\quad 和\quad \mathbf{M}_{R_1\cap R_2}=\mathbf{M}_{R_1}\wedge\mathbf{M}_{R_2}$$

例 4　假设集合 A 上的关系 R_1 和 R_2 由下述矩阵表示

$$\mathbf{M}_{R_1}=\begin{bmatrix}1&0&1\\1&0&0\\0&1&0\end{bmatrix}\quad 和\quad \mathbf{M}_{R_2}=\begin{bmatrix}1&0&1\\0&1&1\\1&0&0\end{bmatrix}$$

什么是 $R_1\cup R_2$ 和 $R_1\cap R_2$ 的矩阵表示？

解　这两个关系的矩阵是

$$\mathbf{M}_{R_1\cup R_2}=\mathbf{M}_{R_1}\vee\mathbf{M}_{R_2}=\begin{bmatrix}1&0&1\\1&1&1\\1&1&0\end{bmatrix}$$

$$\mathbf{M}_{R_1\cap R_2}=\mathbf{M}_{R_1}\wedge\mathbf{M}_{R_2}=\begin{bmatrix}1&0&1\\0&0&0\\0&0&0\end{bmatrix}$$ ■

现在我们来考虑怎样确定关系合成的矩阵。这个矩阵可以通过关系矩阵的布尔积(见 2.7 节)得到。特别地，假设 R 是从 A 到 B 的关系且 S 是从 B 到 C 的关系。又假设 A、B 和 C 分别有 m，n 和 p 个元素。令关于 $S\circ R$，R 和 S 的 0-1 矩阵分别为 $\mathbf{M}_{S\circ R}=[t_{ij}]$，$\mathbf{M}_R=[r_{ij}]$，$\mathbf{M}_S=[s_{ij}]$(这些矩阵的大小分别为 $m\times p$，$m\times n$ 和 $n\times p$)有序对(a_i, c_j)属于 $S\circ R$ 当且仅当存在元素 b_k 使得(a_i, b_k)属于 R 并且(b_k, c_j)属于 S。从而 $t_{ij}=1$，当且仅当对某个 k 有 $r_{ik}=s_{kj}=1$。由布尔积的定义，这意味着

$$\mathbf{M}_{S\circ R}=\mathbf{M}_R\odot\mathbf{M}_S$$

例 5　找出关系 $S\circ R$ 的矩阵，其中表示 R 和 S 的矩阵是

$$\mathbf{M}_R=\begin{bmatrix}1&0&1\\1&1&0\\0&0&0\end{bmatrix}\quad 和\quad \mathbf{M}_S=\begin{bmatrix}0&1&0\\0&0&1\\1&0&1\end{bmatrix}$$

解　关于 $S\circ R$ 的矩阵是

$$\mathbf{M}_{S\circ R}=\mathbf{M}_R\odot\mathbf{M}_S=\begin{bmatrix}1&1&1\\0&1&1\\0&0&0\end{bmatrix}$$ ■

表示两个关系合成的矩阵可以用来求 $\mathbf{M}_{R^n}$ 的矩阵，特别地，由布尔幂的定义有

$$\mathbf{M}_{R^n}=\mathbf{M}_R^{[n]}$$

节末的练习 35 要求证明这个公式。

例 6　求关系 R^2 的矩阵表示，其中 R 的矩阵表示是

$$\mathbf{M}_R=\begin{bmatrix}0&1&0\\0&1&1\\1&0&0\end{bmatrix}$$

解　R^2 的矩阵是

$$\mathbf{M}_{R^2}=\mathbf{M}_R^{[2]}=\begin{bmatrix}0&1&1\\1&1&1\\0&1&0\end{bmatrix}$$ ■

7.3.3　用图表示关系

我们已经显示了一个关系可以通过列出它所有的有序对或使用 0-1 矩阵来表示。还有一种重要的表示关系的方法就是图。把集合的每个元素表示成一个点，每个有序对表示成一条弧，弧上

的箭头标明了弧的方向。当我们把一个有穷集上的关系看作一个有向图时，就可以使用这种图形表示。

定义 1 一个有向图由顶点(或结点)集 V 和边(或弧)集 E 组成，其中边集是 V 中元素的有序对的集合。顶点 a 叫做边(a, b)的始点，而顶点 b 叫做这条边的终点。

形如(a, a)的边用一条从顶点 a 到自身的弧表示。这种边叫做环。

例 7 具有顶点 a，b，c 和 d，边(a, b)，(a, d)，(b, b)，(b, d)，(c, a)，(c, b)和(d, b)的有向图给在图 7-5 中。■

集合 A 上的关系 R 表示成一个有向图，这个图以 A 的元素作为顶点，以有序对(a, b)作为边，其中$(a, b)\in R$。这就在集合 A 上的关系和以 A 作为顶点集的有向图之间构成了一一对应。于是，每一个关于关系的论述对应着一个关于有向图的论述，反之亦然。有向图给出了一个关于关系信息的可见显示。因此，也常常用图研究关系及其性质。(注意从集合 A 到集合 B 的关系可以用一个有向的图示来表示，其中对 A 中的每个元素存在一个顶点，B 中的每个元素存在一个顶点，如 7.1 节所示。然而，当 $A=B$ 时，那种图示表示与这里描述的有向图表示不同，它对关系信息的显示要少得多。)例 8～例 10 说明了怎样用有向图来表示关系。

例 8 集合$\{1, 2, 3, 4\}$上的关系

$$R = \{(1,1),(1,3),(2,1),(2,3),(2,4),(3,1),(3,2),(4,1)\}$$

的有向图显示在图 7-6 中。■

例 9 图 7-7 中的有向图所表示的关系 R 中的有序对是什么？

解 关系中的有序对(x, y)是

$$R = \{(1,3),(1,4),(2,1),(2,2),(2,3),(3,1),(3,3),(4,1),(4,3)\}$$

每个有序对对应了有向图的一条边，(2, 2)和(3, 3)对应了环。■

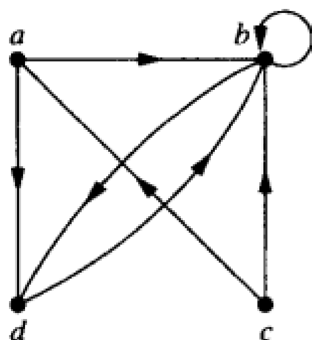


图 7-5 一个有向图

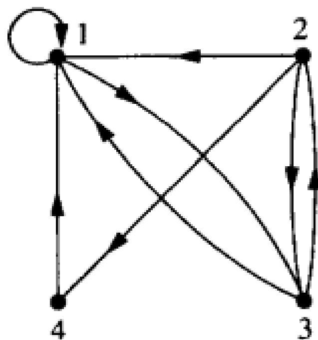


图 7-6 关系 R 的有向图

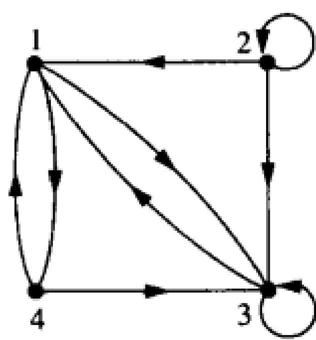
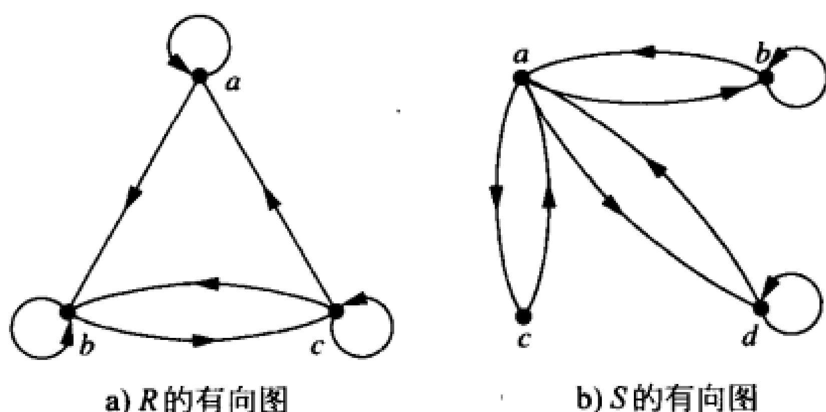


图 7-7 关系 R 的有向图

表示关系的有向图可以用来确定关系是否有各种性质。例如，一个关系是自反的，当且仅当有向图的每个顶点都有环，从而每个形如(x, x)的有序对都出现在关系中。一个关系是对称的，当且仅当对有向图不同顶点之间的每一条边都存在一条方向相反的边，从而只要(x, y)在关系中就有(y, x)在关系中。类似地，一个关系是反对称的，当且仅当在不同的两个顶点之间不存在两条方向相反的边。最后，一个关系是传递的，当且仅当只要存在一条从顶点 x 到顶点 y 的边和一条顶点 y 到顶点 z 的边，就有一条从顶点 x 到顶点 z 的边(完成一个三角形，其中每条边都是具有正确方向的有向边)。

注意 对称关系可以用无向图表示，这个图中的边没有方向。我们将在第 8 章研究无向图。

例 10 对于图 7-8 所示的有向图的关系，确定它们是否为自反的、对称的、反对称的和传递的。



a) R的有向图　　b) S的有向图

图 7-8　关系 R 和 S 的有向图

解　因为 R 的有向图的每个顶点都有环，因此它是自反的。R 既不是对称的也不是反对称的，因为存在一条从 a 到 b 的边，但没有从 b 到 a 的边，并且在连接 b 和 c 的两个方向都有边。最后，R 不是传递的，因为从 a 到 b 有边，从 b 到 c 有边，但是从 a 到 c 没有边。

因为在有向图 S 的所有顶点没都出现环，因此关系不是自反的。因为在不同顶点之间的每条边都伴随一条方向相反的边，因此它是对称的但不是反对称的。不难看出 S 不是传递的，因为(c, a)和(a, b)属于 S 但(c, b)不属于 S。■

练习

1. 用矩阵表示下面每个$\{1, 2, 3\}$上的关系(按增序列出这个集合的元素)。
   a) $\{(1, 1), (1, 2), (1, 3)\}$
   b) $\{(1, 2), (2, 1), (2, 2), (3, 3)\}$
   c) $\{(1, 1), (1, 2), (1, 3), (2, 2), (2, 3), (3, 3)\}$
   d) $\{(1, 3), (3, 1)\}$
2. 用矩阵表示下面每个$\{1, 2, 3, 4\}$上的关系(按增序列出这个集合的元素)。
   a) $\{(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)\}$
   b) $\{(1, 1), (1, 4), (2, 2), (3, 3), (4, 1)\}$
   c) $\{(1, 2), (1, 3), (1, 4), (2, 1), (2, 3), (2, 4), (3, 1), (3, 2), (3, 4), (4, 1), (4, 2), (4, 3)\}$
   d) $\{(2, 4), (3, 1), (3, 2), (3, 4)\}$
3. 列出和下面矩阵对应的$\{1, 2, 3\}$上的关系中的有序对(其中行和列对应于按增序列出的整数)。

   a) $\begin{bmatrix}1 & 0 & 1\\0 & 1 & 0\\1 & 0 & 1\end{bmatrix}$　b) $\begin{bmatrix}0 & 1 & 0\\0 & 1 & 0\\0 & 1 & 0\end{bmatrix}$　c) $\begin{bmatrix}1 & 1 & 1\\1 & 0 & 1\\1 & 1 & 1\end{bmatrix}$

4. 列出和下面矩阵对应的$\{1, 2, 3, 4\}$上的关系中的有序对(其中行和列对应于按增序列出的整数)。

   a) $\begin{bmatrix}1 & 1 & 0 & 1\\1 & 0 & 1 & 0\\0 & 1 & 1 & 1\\1 & 0 & 1 & 1\end{bmatrix}$　b) $\begin{bmatrix}1 & 1 & 1 & 0\\0 & 1 & 0 & 0\\0 & 0 & 1 & 1\\1 & 0 & 0 & 1\end{bmatrix}$　c) $\begin{bmatrix}0 & 1 & 0 & 1\\1 & 0 & 1 & 0\\0 & 1 & 0 & 1\\1 & 0 & 1 & 0\end{bmatrix}$

5. 怎样用表示集合 A 上关系 R 的有向图确定这个关系是否是反自反的?
6. 怎样用表示集合 A 上关系 R 的有向图确定这个关系是否是非对称的?
7. 确定练习 3 中的矩阵所表示的关系是否为自反的、反自反的、对称的、反对称的和传递的。
8. 确定练习 4 中的矩阵所表示的关系是否为自反的、反自反的、对称的、反对称的和传递的。
9. R 是包含了前 100 个正整数的集合 $A=\{1, 2, \cdots, 100\}$上的关系，如果 R 满足下述条件，那么表示 R

的矩阵有多少个非 0 的元素？

a) $\{(a, b) \mid a>b\}$　　b) $\{(a, b) \mid a\neq b\}$　　c) $\{(a, b) \mid a=b+1\}$

d) $\{(a, b) \mid a=0\}$　　e) $\{(a, b) \mid ab=0\}$

10. R 是包含了前 1000 个正整数的集合 $A=\{1, 2, \cdots, 1000\}$ 上的关系，如果 R 满足下述条件，那么表示 R 的矩阵有多少个非 0 的元素？

a) $\{(a, b) \mid a\leqslant b\}$　　b) $\{(a, b) \mid a=b\pm 1\}$　　c) $\{(a, b) \mid a+b=1000\}$

d) $\{(a, b) \mid a+b\leqslant 1001\}$　　e) $\{(a, b) \mid a\neq 0\}$

11. 当 R 是有穷集 A 上的关系时，怎样从表示 R 的关系矩阵找到这个关系的补 $\overline{R}$ 的矩阵？

12. 当 R 是有穷集 A 上的关系时，怎样从表示 R 的关系矩阵找到这个关系的逆 R^{-1} 的矩阵？

13. 设 R 是矩阵

$$\mathbf{M}_R=\begin{bmatrix}0 & 1 & 1\\ 1 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}$$

所表示的关系，求表示下述关系的矩阵。

a) R^{-1}　　b) $\overline{R}$　　c) R^2

14. 设 R_1 和 R_2 是集合 A 上的关系并由矩阵

$$\mathbf{M}_{R_1}=\begin{bmatrix}0 & 1 & 0\\ 1 & 1 & 1\\ 1 & 0 & 0\end{bmatrix}\quad 和\quad \mathbf{M}_{R_2}=\begin{bmatrix}0 & 1 & 0\\ 0 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}$$

表示。求表示下述关系的矩阵。

a) $R_1\cup R_2$　　b) $R_1\cap R_2$　　c) $R_2\circ R_1$

d) $R_1\circ R_1$　　e) $R_1\oplus R_2$

15. 设 R 是矩阵

$\mathbf{M}_R=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 1 & 0\end{bmatrix}$表示的关系，求表示下述关系的矩阵

a) R^2　　b) R^3　　c) R^4

16. 设 R 是 n 元素集合上的关系。如果在表示 R 的矩阵 $\mathbf{M}_R$ 中存在 k 个非 0 的元素，那么在表示 R 的逆 R^{-1} 的矩阵 $\mathbf{M}_{R^{-1}}$ 中存在多少个非 0 元素？

17. 设 R 是 n 元素集合上的关系。如果在表示 R 的矩阵 $\mathbf{M}_R$ 中存在 k 个非 0 的元素，那么在表示 R 的补 $\overline{R}$ 的矩阵 $\mathbf{M}_{\overline{R}}$ 中存在多少个非 0 元素？

18. 画出表示练习 1 中每个关系的有向图。

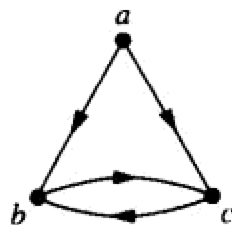
19. 画出表示练习 2 中每个关系的有向图。

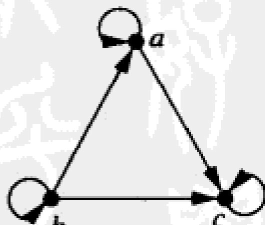
20. 画出表示练习 3 中每个关系的有向图。

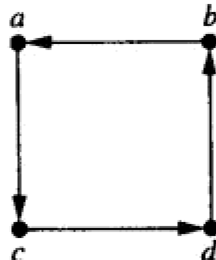
21. 画出表示练习 4 中每个关系的有向图。

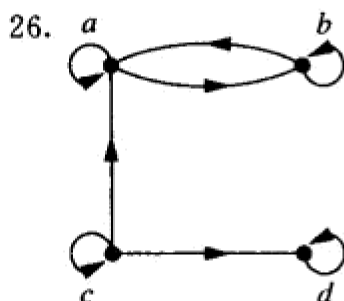
22. 画出表示关系 $\{(a, a), (a, b), (b, c), (c, b), (c, d), (d, a), (d, b)\}$ 的有向图。

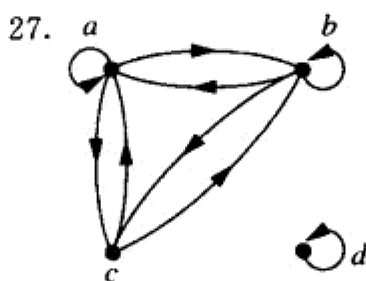
在练习 23～28 列出由下述有向图所表示的关系中的有序对。

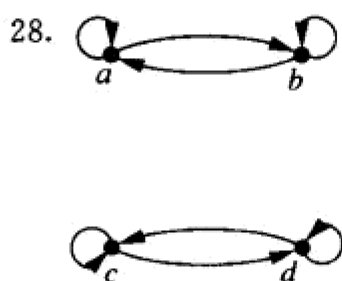
23.


24.


25.


26. 

27. 

28. 

29. 怎样用有穷集 A 上关系 R 的有向图确定关系是否是非对称的?
30. 怎样用有穷集 A 上关系 R 的有向图确定关系是否是反自反的?
31. 确定练习 23～25 所示的有向图表示的关系是否为自反的、反自反的、对称的、反对称的和传递的。
32. 确定练习 26～28 所示的有向图表示的关系是否为自反的、反自反的、对称的、反对称的和传递的。
33. 设 R 是集合 A 上的关系，解释怎样用表示 R 的有向图得到表示关系的逆 R^{-1} 的有向图?
34. 设 R 是集合 A 上的关系，解释怎样用表示 R 的有向图得到表示关系的补 $\bar{R}$ 的有向图?
35. 证明如果 $\mathbf{M}_R$ 是表示关系 R 的矩阵，那么 $\mathbf{M}_R^{[n]}$ 是表示关系 R^n 的矩阵。
36. 给定表示两个关系的有向图，怎样找出这些关系的并、交、对称差、差以及合成的有向图?

7.4 关系的闭包

7.4.1 引言

一个计算机网络在波士顿、芝加哥、丹佛、底特律、纽约和圣迭戈设有数据中心。从波士顿到芝加哥，波士顿到底特律，芝加哥到底特律，底特律到丹佛和纽约到圣迭戈都有单向的电话线。如果存在一条从数据中心 a 到 b 的电话线，(a, b)就属于关系 R。我们怎样确定从一个中心是否有一条电话线或多条电话线(可能不直接)链接到另一个中心? 由于所有的链接不一定是直接的，例如从波士顿可通过底特律到丹佛，因此不能直接使用 R 来回答这个问题。用关系的语言说，R 不是传递的，因而它不包含可能被链接的所有的对。正如我们将在本节证明的，我们可以通过构造包含 R 的最小的传递关系来找出每一对有着链接的数据中心。这个关系叫做 R 的传递闭包。

一般说来，设 R 是集合 A 上的关系。R 可能具有或者不具有某些性质 **P**，例如自反性、对称性或传递性。如果存在包含 R 的具有性质 **P** 的关系 S，并且 S 是包含 R 且具有性质 **P** 的每一个关系的子集，那么 S 叫做 R 的关于 **P** 的闭包。(注意一个关系关于一个性质的闭包可能不存在，见节末的练习 15 和 35。)我们将说明怎样找关系的自反闭包、对称闭包和传递闭包。

7.4.2 闭包

集合 $A=\{1, 2, 3\}$上的关系 $R=\{(1, 1), (1, 2), (2, 1), (3, 2)\}$不是自反的。我们怎样才能产生一个包含 R 的尽可能小的自反关系呢? 这可以通过把$(2, 2)$和$(3, 3)$加到 R 中来做到，因为只有它们是不在 R 中的形如(a, a)的对。很明显，这个新关系包含了 R。此外，任何包含 R 的自反关系必须包含$(2, 2)$和$(3, 3)$。因为这个关系包含了 R，是自反的，并且被包含在每一个包含 R 的自反关系之中，因此它就是 R 的自反闭包。

正如这个例子所显示的，给定集合 A 上的关系 R，对于 $a\in R$，可以通过把形如(a, a)的所有的对，除了已在 R 中的之外，都加到 R 中，就构成了 R 的自反闭包。加入这些对产生了一个新的自反的、包含 R 的关系，并且它被包含在任何包含 R 的自反关系之中。我们看到 R 的自反闭包等于 $R\cup\Delta$，其中 $\Delta=\{(a, a) \mid a\in A\}$是 A 上的对角关系。(读者应能验证这个结果。)

例 1　整数集上的关系 $R=\{(a, b) \mid a<b\}$的自反闭包是什么?

解 R的自反闭包是

$$R \cup \Delta = \{(a,b) \mid a < b\} \cup \{(a,a) \mid a \in \mathbf{Z}\} = \{(a,b) \mid a \leqslant b\}$$ ■

{1，2，3}上的关系{(1，1)，(1，2)，(2，2)，(2，3)，(3，1)，(3，2)}不是对称的。我们怎样产生一个包含R的尽可能小的对称关系？只需增加(2，1)和(1，3)，因为只有它们是具有$(a, b)\in R$而(b, a)不在R中的那种(b, a)对。这个新关系是对称的，且包含了R。此外，任何包含了R的对称关系一定包含这个新关系，因为一个包含R的对称关系必须包含(2，1)和(1，3)。因此这个新关系叫做R的对称闭包。

正如这个例子所显示的，关系R的对称闭包可以通过增加所有形如(b, a)的有序对构成，其中(a, b)在关系中而(b, a)不在关系中。增加这些有序对产生一个关系，它是对称的，包含了R，并且它被包含在任何包含R的对称关系之中。关系R的对称闭包可以通过取关系与它的逆的并来构造，即$R\cup R^{-1}$是R的对称闭包，其中$R^{-1}=\{(b, a) \mid (a, b)\in R\}$。读者应能验证这个结果。

例2 正整数集合上的关系$R=\{(a, b) \mid a>b\}$的对称闭包是什么？

解 R的对称闭包是关系

$$R \cup R^{-1} = \{(a,b) \mid a > b\} \cup \{(b,a) \mid a > b\} = \{(a,b) \mid a \neq b\}$$

这最后一步相等是由于R包含了所有正整数构成的有序对，其中第一元素大于第二元素，并且R^{-1}包含了所有正整数构成的有序对，其中第一元素小于第二元素。 ■

假设关系R不是传递的，我们怎样产生一个包含R的传递关系并使得这个新的关系被包含在任何包含R的传递关系之中？对于已经在R中的(a, b)和(b, c)，可以通过增加所有形如(a, c)的有序对构成R的传递闭包吗？考虑集合{1，2，3，4}上的关系R={(1，3)，(1，4)，(2，1)，(3，2)}。这个关系不是传递的，因为对于在R中的(a, b)和(b, c)，它不包含所有形如(a, c)的有序对。这种不在R中的有序对是(1，2)，(2，3)，(2，4)和(3，1)。把这些有序对加到R中并不能产生一个传递关系，因为所得的结果关系包含(3，1)和(1，4)，但不包含(3，4)。这说明构造关系的传递闭包比起构造它们的自反闭包或对称闭包更复杂。下面我们将建立一个构造传递闭包的算法。正如将要显示的，一个关系的传递闭包可以通过增加那些必须出现的有序对来得到，要重复这个过程，直到没有必须增加的有序对为止。

7.4.3 有向图的路径

我们将看到用有向图表示关系有助于构造关系的传递闭包。为此现在引入某些将要用到的术语。

通过沿有向图的边(按照这条边的箭头指示的相同方向)移经有向图就得到一条有向图中的路径。

定义1 在有向图G中从a到b的一条路径是G中一条或多条边的序列(x_0, x_1)，(x_1, x_2)，(x_2, x_3)，…，(x_{n-1}, x_n)，其中$x_0=a$，$x_n=b$。即一个边的序列，其中一条边的终点和路径中下一条边的始点相同。这条路径记为x_0，x_1，…，x_{n-1}，x_n，长度为n。在同一顶点开始和结束的路径叫做**回路**或**圈**。

有向图的一条路径可以多次通过一个顶点。此外，有向图的一条边也可以多次出现在一条路径中。

例3 下面哪些路径是图7-9的有向图中的路径：a，b，e，d；a，e，c，d，b；b，a，c，b，a，a，b；d，c；c，b，a；e，b，a，b，a，b，e？这些路径的长度是多少？这个列表中的哪些路径是回路？

解　因为(a, b)、(b, e)和(e, d)都是边，a, b, e, d是长为3的路径。因为(c, d)不是边，a, e, c, d, b不是路径。因为(b, a)，(a, c)，(c, b)，(b, a)，(a, a)和(a, b)都是边，b, a, c, b, a, a, b是长为6的路径。我们也看到，因为(d, c)是边，d, c是长为1的路径。还有由于(c, b)和(b, a)是边，c, b, a是长为2的路径。(e, b)，(b, a)，(a, b)，(b, a)，(a, b)，(b, e)都是边，因此e, b, a, b, a, b, e是长为6的路径。

两条路径b, a, c, b, a, a, b和e, b, a, b, a, b, e是回路，因为它们在同一顶点开始和结束。路径a, b, e, d；c, b, a和d, c不是回路。■

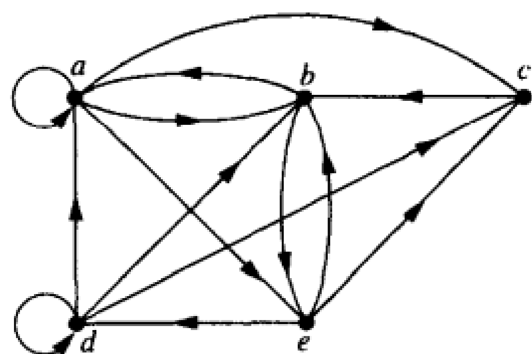


图 7-9　一个有向图

术语路径也用于关系。把有向图的定义推广到关系可知，如果存在一个元素的序列$a, x_1, x_2, \cdots, x_{n-1}, b$具有$(a, x_1) \in R$，$(x_1, x_2) \in R$，$\cdots$，$(x_{n-1}, b) \in R$，那么在$R$中存在一条从$a$到$b$的路径。从关系中的路径定义可以得到定理1。

定理 1　设R是集合A上的关系。从a到b存在一条长为n(n为正整数)的路径，当且仅当$(a, b) \in R^n$。

证　我们将使用数学归纳法。根据定义，从a到b存在一条长为1的路径，当且仅当$(a, b) \in R$。因此当$n=1$时定理为真。

假定对于正整数n定理为真，这是归纳假设。从a到b存在一条长为$n+1$的路径，当且仅当存在元素$c \in A$使得从a到c存在一条长为1的路径，即$(a, c) \in R$，以及一条从c到b的长为n的路径，即$(c, b) \in R^n$。因此，由归纳假设，从a到b存在一条长为$n+1$的路径，当且仅当存在一个元素c，使得$(a, c) \in R$和$(c, b) \in R^n$。但是存在这样一个元素，当且仅当$(a, b) \in R^{n+1}$。因此，从a到b存在一条长为$n+1$的路径，当且仅当$(a, b) \in R^{n+1}$。定理得证。□

7.4.4　传递闭包

我们现在证明找一个关系的传递闭包与在相关的有向图中确定哪些顶点对被路径连接是等价的。由此要定义一个新的关系。

定义 2　设R是集合A上的关系。连通性关系R^*由对(a, b)构成，使得在R中从顶点a到b之间存在一条至少长为1的路径。

因为R^n由对(a, b)构成，使得存在一条从a到b的长为n的路径，从而R^*是所有集合R^n的并。换句话说，

$$R^* = \bigcup_{n=1}^{\infty} R^n$$

许多模型都用到连通性关系。

例 4　设R是世界上所有人的集合上的关系，如果a认识b，那么R包含(a, b)。R^n是什么？其中n是大于2的正整数。R^*是什么？

解　如果存在c使得$(a, c) \in R$且$(c, b) \in R$，即存在c使得a认识c，c认识b，那么关系R^2包含(a, b)。类似地，如存在$x_1, x_2, \cdots, x_{n-1}$使得$a$认识$x_1$，$x_1$认识$x_2$，$\cdots$，$x_{n-1}$认识$b$，那么$R^n$包含对$(a, b)$。

如果存在从a开始至b终止的序列，使得序列中的每个人都认识序列中的下一个人，那么R^*包含对(a, b)。(关于R^*存在许多有趣的猜想。你认为这个连通关系包含以你作为第一元素，蒙古的总统作为第二元素的一对元素吗？在第8章我们将用图建立这个应用的模型。)■

例 5　设R是纽约市所有地铁站集合的关系。如果可以从a站不换车就旅行到b站，那么R

包含对(a, b)。当 n 是正整数时，R^n 是什么？R^* 是什么？

解 如果换 $n-1$ 次车就可以从 a 站旅行到 b 站，关系 R^n 就包含(a, b)。从 a 站旅行到 b 站，如果需要可以换车任意多次，关系 R^* 就由这种有序对(a, b)组成。(读者应该能验证这些论断。) ■

例 6 设 R 是美国所有的州的集合上的关系，如果 a 州和 b 州有公共的边界，那么 R 包含(a, b)。R^n 是什么？其中 n 是正整数。R^* 又是什么？

解 关系 R^n 由对(a, b)构成，可以从 a 州恰好跨越 n 次州界到达 b 州。R^* 由对(a, b)构成，可以从 a 州跨越任意多次边界到达 b 州(读者应该能够验证这些论断。)那些包含与美国大陆不相连的州(即含有阿拉斯加或夏威夷)的有序对是不在 R^* 中的。 ■

定理 2 证明一个关系的传递闭包和相关的连通性关系是等同的。

定理 2 关系 R 的传递闭包等于连通性关系 R^*。

证 注意 R^* 包含 R。为证明 R^* 是 R 的传递闭包，我们必须证明 R^* 是传递的且对一切包含 R 的传递关系 S，有 $R^* \subseteq S$。

首先，我们证明 R^* 是传递的。如果$(a, b) \in R^*$ 和$(b, c) \in R^*$，那么在 R 中存在从 a 到 b 和从 b 到 c 的路径。我们从 a 到 b 的路径开始，并且沿着从 b 到 c 的路径就得到一条从 a 到 c 的路径。因此，$(a, c) \in R^*$。这就得出 R^* 是传递的。

现在假设 S 是包含 R 的传递关系，因为 S 是传递的，S^n 也是传递的(读者应该能验证这一点)，并且 $S^n \subseteq S$(由 7.1 节定理 1)。此外，因为

$$S^* = \bigcup_{k=1}^{\infty} S^k$$

和 $S^k \subseteq S$，因此 $S^* \subseteq S$。现在注意到如果 $R \subseteq S$，那么 $R^* \subseteq S^*$，这是由于任何 R 中的路径也是 S 中的路径。从而 $R^* \subseteq S^* \subseteq S$。于是，任何包含 R 的传递关系也一定包含 R^*。因此，R^* 是 R 的传递闭包。 □

既然我们知道传递闭包等于连通性关系，我们考虑这个关系的计算问题。在一个有限的有向图中不需要检测任意长的路径来确定是否在两个顶点之间存在一条路径。正如下面的引理所证明的。检测包含不超过 n 条边的路径就足够了，这里 n 是集合中的元素个数。

引理 1 设 A 是 n 元素集合，R 是 A 上的关系。如果 R 中存在一条从 a 到 b 的长至少为 1 的路径，那么存在一条长度不超过 n 的这种路径。此外，当 $a \neq b$ 时，如果在 R 中存在一条从 a 到 b 的路径，那么存在一条长度不超过 $n-1$ 的这种路径。

证 假设 R 中存在一条从 a 到 b 的路径。令 m 是这种路径的最短长度。假设 $x_0, x_1, x_2, \cdots, x_{m-1}, x_m$ 是一条这样的路径，其中 $x_0 = a$，$x_m = b$。

假设 $a = b$ 和 $m > n$，使得 $m \geq n+1$。由鸽巢原理，因为 A 中有 n 个顶点，在 m 个顶点 $x_0, x_1, x_2, \cdots, x_{m-1}$ 中至少有两个是相同的(见图 7-10)。

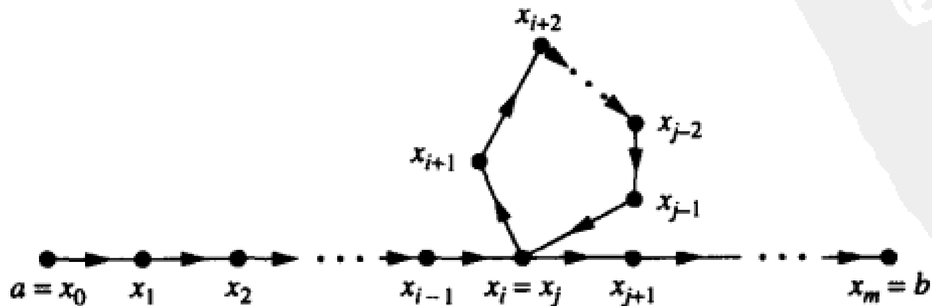


图 7-10 产生一条长度不超过 n 的路径

假设 $x_i=x_j$ 满足 $0\leqslant i<j\leqslant m-1$。那么这条路径包含一条从 x_i 到 x_i 自身的回路。可以把这条回路从由 a 到 b 的路径中删除，剩下的路径，即 x_0，x_1，…，x_i，x_{j+1}，…，x_{m-1}，x_m 是从 a 到 b 的更短的路径。因此，最短长度的这种路径的长度一定小于等于 n。

$a\neq b$ 的情况留给读者作为练习。 □

由引理 1，我们看出 R 的传递闭包是 R，R^2，R^3，…，R^n 的并。这是由于在 R^* 的两个顶点之间存在一条路径，当且仅当对某个正整数 $i(i\leqslant n)$ 在 R^i 的这些顶点之间存在一条路径。因为

$$R^* = R \cup R^2 \cup R^3 \cup \cdots \cup R^n$$

并且表示关系的并的 0-1 矩阵是这些关系的 0-1 矩阵的联合，因此传递闭包的 0-1 矩阵是 R 的 0-1 矩阵的前 n 次幂的 0-1 矩阵的联合。

定理 3 设 $\mathbf{M}_R$ 是 n 元素集合上的关系 R 的 0-1 矩阵。那么传递闭包 R^* 的 0-1 矩阵是

$$\mathbf{M}_{R^*} = \mathbf{M}_R \vee \mathbf{M}_R^{[2]} \vee \mathbf{M}_R^{[3]} \vee \cdots \vee \mathbf{M}_R^{[n]}$$

例 7 求关系 R 的传递闭包的 0-1 矩阵，其中

$$\mathbf{M}_R = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}$$

解 由定理 3，R^* 的 0-1 矩阵是

$$\mathbf{M}_{R^*} = \mathbf{M}_R \vee \mathbf{M}_R^{[2]} \vee \mathbf{M}_R^{[3]}$$

因为

$$\mathbf{M}_R^{[2]} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{和} \quad \mathbf{M}_R^{[3]} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

所以

$$\mathbf{M}_{R^*} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \vee \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \vee \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$ ■

定理 3 可以作为计算关系 R^* 的矩阵的算法基础。为求出这个矩阵，要连续计算 $\mathbf{M}_R$ 的布尔幂，直到第 n 次幂为止。当计算每次幂时就构成这个幂与所有较小的幂的联合。当做到第 n 次幂时，就得到关于 R^* 的矩阵。这个过程见算法 1。

算法 1 计算传递闭包的过程

```
procedure transitive closure (M_R: n×n 的 0-1 矩阵)
A := M_R
B := A
for i := 2 to n
begin
  A := A ⊙ M_R
  B := B ∨ A
end{B 是关于 R* 的 0-1 矩阵}
```

我们可以容易地求出用算法 1 确定关系的传递闭包所使用的位运算次数。计算布尔幂 $\mathbf{M}_R$，$\mathbf{M}_R^{[2]}$，…，$\mathbf{M}_R^{[n]}$ 需要找到 $n-1$ 个 $n\times n$ 的 0-1 矩阵的布尔积。每个布尔积可以使用 $n^2(2n-1)$ 次

位运算求得。因此，计算这些乘积使用 $n^2(2n-1)(n-1)$ 次位运算。

为从 n 个 $\mathbf{M}_R$ 的布尔幂求 $\mathbf{M}_{R^*}$，需要求 $n-1$ 个 0-1 矩阵的联合。计算每一个联合使用 n^2 次位运算。因此，在这部分计算中使用 $(n-1)n^2$ 次位运算。所以，当使用算法 1 时，用 $n^2(2n-1)(n-1)+(n-1)n^2=2n^3(n-1)=O(n^4)$ 次位运算就可以求出 n 元素集合上关系的传递闭包的矩阵。下面我们将要描述一个更有效的求传递闭包的算法。

7.4.5 沃舍尔算法

沃舍尔算法得名于史蒂芬·沃舍尔[⊖]，他在 1960 年给出该算法。这个算法能够有效计算关系的传递闭包。算法 1 求出 n 元素集合上关系的传递闭包使用 $2n^3(n-1)$ 次位运算。而沃舍尔算法只使用 $2n^3$ 次位运算就可以求出这个传递闭包。

注意 *沃舍尔算法有时也叫做罗伊-沃舍尔算法，因为伯纳德·罗伊(B·Roy)在 1959 年描述了这个算法。*

假设 R 是 n 元素集合上的关系，设 $v_1, v_2, \cdots, v_n$ 是这 n 个元素的任意排列。沃舍尔算法中用到一条路径的内点的概念。如果 $a, x_1, x_2, \cdots, x_{m-1}, b$ 是一条路径，它的内点是 $x_1, x_2, \cdots, x_{m-1}$，即除了第一和最后一个顶点之外出现在路径中的所有顶点。例如，在有向图的一条路径 a, c, d, f, g, h, b, j 的内点是 c, d, f, g, h, b。a, c, d, a, f, b 的内点是 c, d, a, f。(注意这条路径的起点不是内点，除非这条路径再次访问它，且不是作为终点来访问的。类似地，这条路径的终点也不是内点，除非它在这之前曾被这条路径访问过，且不是作为起点来访问的。)

沃舍尔算法的基础是构造一系列 0-1 矩阵。这些矩阵是 $\mathbf{W}_0, \mathbf{W}_1, \cdots, \mathbf{W}_n$，其中 $\mathbf{W}_0=\mathbf{M}_R$ 是这个关系的 0-1 矩阵，且 $\mathbf{W}_k=[w_{ij}^{[k]}]$。如果存在一条从 v_i 到 v_j 的路径使得这条路径的所有内点都在集合 $\{v_1, v_2, \cdots, v_k\}$(表中的前 k 个顶点)之中，那么 $w_{ij}^{[k]}=1$，否则为 0(这条路径的起点和终点可能在表中的前 k 个顶点的集合之外)。注意 $\mathbf{W}_n=\mathbf{M}_{R^*}$，因为 $\mathbf{M}_{R^*}$ 的第 (i, j) 项是 1，当且仅当存在一条从 v_i 到 v_j 的路径，且全部内点都在集合 $\{v_1, v_2, \cdots, v_n\}$ 之中(但这些就是有向图的所有顶点)。例 8 说明矩阵 $\mathbf{W}_k$ 表示什么。

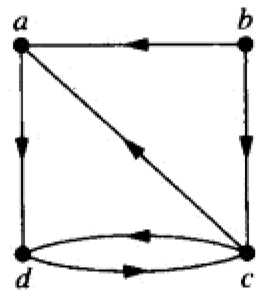


图 7-11 关系 R 的有向图

例 8 设 R 是一个关系，它的有向图如图 7-11 所示。设 a, b, c, d 是集合元素的排列。求矩阵 $\mathbf{W}_0, \mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3$ 和 $\mathbf{W}_4$。矩阵 $\mathbf{W}_4$ 是 R 的传递闭包。

⊖ **史蒂芬·沃舍尔(Stephen Warshall，生于 1935 年)** 沃舍尔生于纽约，在布鲁克莱恩的公立学校接受教育。他 1956 年于哈佛大学获得数学学位。他没有得到更高的学位，因为在那个年代在他感兴趣的领域没有合适的培养计划。但是，他在几个不同的大学研修了研究生课程，并且对计算机科学和软件工程的发展作出了贡献。

从哈佛大学毕业以后，沃舍尔在 ORO(运筹学办公室)工作，它由约翰·霍普金斯建立，为美国陆军做研究和开发工作。1958 年他离开 ORO 去，一个叫做 Technical Operations 的公司工作。在那里他帮助建立了一个从事军事软件课题研究和开发的实验室。1961 年他离开了这家公司，创建马萨诸塞计算机联合公司。后来，这个公司变成了应用数据研究(ADR)公司的一部分。在这次合并以后，沃舍尔进入了 ADR 的董事会，管理各种项目和组织。他 1982 年从 ADR 退休。

在任职期间，沃舍尔在操作系统、编译器设计、语言设计和运筹学领域从事研究和开发工作。在 1971—1972 年学术年会上他在法国的大学作了关于软件工程方面的报告。关于这个传递闭包算法，即目前所称的沃舍尔算法的正确性证明有一个有趣的故事。他和一个在 Technical Operations 公司的同事打赌，首先确定这个算法是永远有效的一方将赢得一瓶甜酒。沃舍尔一夜之后给出了证明，赌赢了，并和这个同事分享了这瓶甜酒。因为沃舍尔不喜欢坐在写字桌旁，他的许多创造性工作都是在印度洋的帆船或希腊的柠檬园这种不寻常的地方完成的。

解　令 $v_1=a$，$v_2=b$，$v_3=c$，且 $v_4=d$。$\mathbf{W}_0$ 是这个关系的矩阵，于是

$$\mathbf{W}_0=\begin{bmatrix}0&0&0&1\\1&0&1&0\\1&0&0&1\\0&0&1&0\end{bmatrix}$$

如果存在一条 v_i 到 v_j 的且只有 $v_1=a$ 作为内点的路径，$\mathbf{W}_1$ 的$(i,\ j)$项有 1。注意因为所有长为 1 的路径没有内点，所以仍旧可以使用这些路径。此外存在一条从 b 到 d 的路径，即 b，a，d。因此

$$\mathbf{W}_1=\begin{bmatrix}0&0&0&1\\1&0&1&1\\1&0&0&1\\0&0&1&0\end{bmatrix}$$

如果存在一条 v_i 到 v_j 的且只有 $v_1=a$ 和 $v_2=b$ 作为内点的路径，$\mathbf{W}_2$ 的$(i,\ j)$项有 1。因为没有边以 b 作为终点，当我们允许 b 作为内点时不会得到新的路径。因此，$\mathbf{W}_2=\mathbf{W}_1$。

若存在一条 v_i 到 v_j 的只用 $v_1=a$，$v_2=b$ 和 $v_3=c$ 作为内点的路径，则 $\mathbf{W}_3$ 的$(i,\ j)$项有 1。我们现在有从 d 到 a 的路径，即 d，c，a 和从 d 到 d 的路径 d，c，d。因此

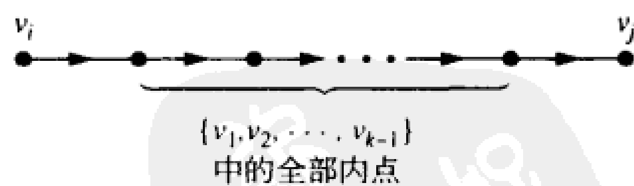
$$\mathbf{W}_3=\begin{bmatrix}0&0&0&1\\1&0&1&1\\1&0&0&1\\1&0&1&1\end{bmatrix}$$

最后，如果存在一条从 v_i 到 v_j 的路径，并且以 $v_1=a$，$v_2=b$，$v_3=c$ 及 $v_4=d$ 作为内点，那么 $\mathbf{W}_4$ 的$(i,\ j)$项为 1。因为这些是图的全部顶点，此项为 1，当且仅当存在一条从 v_i 到 v_j 的路径。因此

$$\mathbf{W}_4=\begin{bmatrix}1&0&1&1\\1&0&1&1\\1&0&1&1\\1&0&1&1\end{bmatrix}$$

这个最后的矩阵就是传递闭包的矩阵。■

沃舍尔算法通过有效地计算 $\mathbf{W}_0=\mathbf{M}_R$，$\mathbf{W}_1$，$\mathbf{W}_2$，…，$\mathbf{W}_n=\mathbf{M}_{R^*}$ 来计算 $\mathbf{M}_{R^*}$。不难看出可以直接从 $\mathbf{W}_{k-1}$ 计算 $\mathbf{W}_k$：存在一条从 v_i 到 v_j 的只以 v_1，v_2，…，v_k 中的顶点作为内点的路径，当且仅当要么存在一条从 v_i 到 v_j 的且内点是表中前 $k-1$ 个顶点的路径，要么存在从 v_i 到 v_k 的路径和从 v_k 到 v_j 的路径，而这些路径的内点仅在表中的前 $k-1$ 个顶点中。这就是说，要么在 v_k 被允许作为内点之前从 v_i 到 v_j 已经存在一条路径，要么允许 v_k 作为内点产生一条从 v_i 到 v_k 然后从 v_k 到 v_j 的路径。这两种情况如图 7-12 所示。



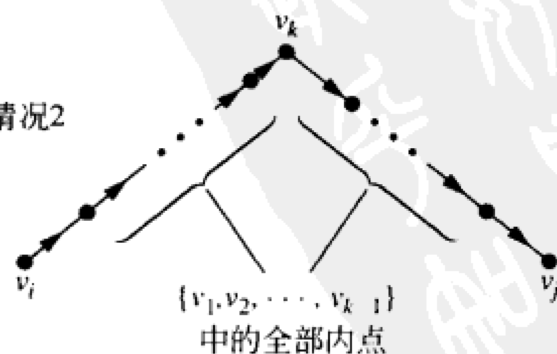


图 7-12　把 v_k 加到允许使用的内点集中

第一种类型的路径存在，当且仅当 $w_{ij}^{[k-1]}=1$，且第二种类型的路径存在，当且仅当 $w_{ik}^{[k-1]}=1$ 和 $w_{kj}^{[k-1]}=1$。于是，$w_{ij}^{[k]}$ 是 1，当且仅当或者 $w_{ij}^{[k-1]}=$

1，或者 $w_{ik}^{[k-1]}=1$ 和 $w_{kj}^{[k-1]}=1$。这就得到引理 2。

引理 2　设 $\mathbf{W}_k=[w_{ij}^{[k]}]$是 0-1 矩阵，它在$(i, j)$位置有 1，当且仅当存在一条从 v_i 到 v_j 的路径，其内点取自集合$\{v_1, v_2, \cdots, v_k\}$，那么

$$w_{ij}^{[k]}=w_{ij}^{[k-1]} \vee (w_{ik}^{[k-1]} \wedge w_{kj}^{[k-1]})$$

其中 i，j 和 k 是不超过 n 的正整数。

引理 2 提供了有效计算矩阵 $\mathbf{W}_k(k=1, 2, \cdots, n)$的手段。我们使用引理 2 把沃舍尔算法的伪码在算法 2 中给出。

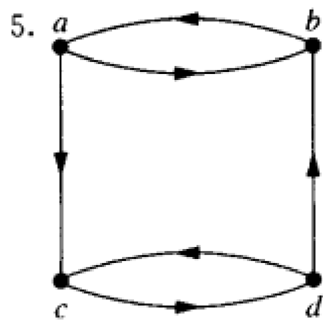
```
算法 2  沃舍尔算法
procedure Warshall(M_R: n×n 的 0-1 矩阵)
W := M_R
for k := 1 to n
begin
    for i := 1 to n
    begin
      for j := 1 to n
        w_ij := w_ij ∨ (w_ik ∧ w_kj)
    end
end{W=[w_ij]是 M_R*}
```

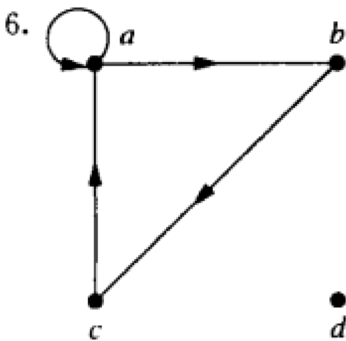
很容易以位运算次数计算沃舍尔算法的计算复杂度。使用引理 2，从项 $w_{ij}^{[k-1]}$，$w_{ik}^{[k-1]}$和 $w_{kj}^{[k-1]}$找到项 $w_{ij}^{[k]}$需要 2 次位运算。为从 $\mathbf{W}_{k-1}$求出 $\mathbf{W}_k$ 的所有 n^2 个项需要 $2n^2$ 次位运算。因为沃舍尔算法从 $\mathbf{W}_0=\mathbf{M}_R$ 开始，计算 n 个 0-1 矩阵的序列 $\mathbf{W}_1$，$\mathbf{W}_2$，…，$\mathbf{W}_n=\mathbf{M}_{R^*}$，使用的位运算总次数是 $n \cdot 2n^2=2n^3$。

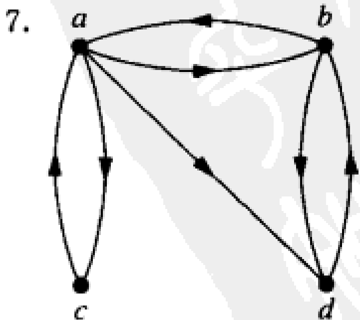
练习

1. 设 R 是集合$\{0, 1, 2, 3\}$上的关系，R 包含有序对$(0, 1)$，$(1, 1)$，$(1, 2)$，$(2, 0)$，$(2, 2)$和$(3, 0)$，求
   a)R 的自反闭包　　　　b)R 的对称闭包
2. 设 R 是整数集上的关系$\{(a, b) \mid a \neq b\}$，R 的自反闭包是什么？
3. 设 R 是整数集上的关系$\{(a, b) \mid a$ 整除 $b\}$，R 的对称闭包是什么？
4. 从有穷集上关系的有向图怎样构造表示它的自反闭包的有向图？

在练习 5～7 中画出给定有向图所表示关系的自反闭包的有向图。

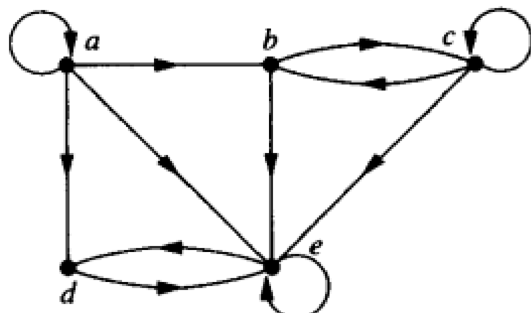
5. 

6. 

7. 

8. 从有穷集上关系的有向图怎样构造表示它的对称闭包的有向图？

9. 对于练习 5～7 的有向图表示的关系，找出关系的对称闭包的有向图。
10. 找出包含了例 2 中关系的最小的自反和对称的关系。
11. 对于练习 5～7 的每个有向图表示的关系，求包含它的最小的自反对称关系的有向图。
12. 假设有穷集 A 上的关系 R 由矩阵 $\mathbf{M}_R$ 表示，证明表示 R 的自反闭包的矩阵是 $\mathbf{M}_R \vee \mathbf{I}_n$。
13. 假设有穷集 A 上的关系 R 由矩阵 $\mathbf{M}_R$ 表示，证明表示 R 的对称闭包的矩阵是 $\mathbf{M}_R \vee \mathbf{M}_R^t$。
14. 证明关系 R 关于性质 $\mathbf{P}$ 的闭包如果存在，就是所有包含 R 的具有性质 $\mathbf{P}$ 的关系的交。
15. 什么时候可能定义一个关系 R 的反自反闭包，即一个包含 R 的关系是反自反的且被包含在每一个包含 R 的反自反关系之中？



16. 确定下面的顶点序列是否为右面的有向图中的路径。

a) a, b, c, e　　b) b, e, c, b, e

c) a, a, b, e, d, e　　d) b, c, e, d, a, a, b

e) b, c, c, b, e, d, e, d　　f) $a, a, b, b, c, c, b, e, d$

17. 求出练习 16 的有向图中所有长为 3 的路径。
18. 确定练习 16 的有向图中是否存在一条以下面给定的第一顶点作为起点、以第二顶点作为终点的路径。

a) a, b　　b) b, a　　c) b, b　　d) a, e　　e) b, d

f) c, d　　g) d, d　　h) e, a　　i) e, c

19. 设 R 是集合{1, 2, 3, 4, 5}上的关系，R 包含有序对(1, 3)，(2, 4)，(3, 1)，(3, 5)，(4, 3)，(5, 1)，(5, 2)和(5, 4)。求

a) R^2　　b) R^3　　c) R^4　　d) R^5　　e) R^6　　f) R^*

20. 设 R 是关系，如果存在一条从 a 城到 b 城的直达航班，则 R 包含有序对(a, b)。什么时候(a, b)在下面的关系中？

a) R^2　　b) R^3　　c) R^*

21. 设 R 是所有学生的集合上的关系，如果 $a \neq b$ 且 a 和 b 至少有一门公共课程，则 R 包含了有序对(a, b)。什么时候(a, b)在下面的关系中？

a) R^2　　b) R^3　　c) R^*

22. 假设关系 R 是自反的，证明 R^* 是自反的。
23. 假设关系 R 是对称的，证明 R^* 是对称的。
24. 假设关系 R 是反自反的，关系 R^2 一定是反自反的吗？
25. 使用算法 1 找出下面{1, 2, 3, 4}上的关系的传递闭包。

a) {(1, 2), (2, 1), (2, 3), (3, 4), (4, 1)}

b) {(2, 1), (2, 3), (3, 1), (3, 4), (4, 1), (4, 3)}

c) {(1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)}

d) {(1, 1), (1, 4), (2, 1), (2, 3), (3, 1), (3, 2), (3, 4), (4, 2)}

26. 使用算法 1 找出下面{a, b, c, d, e}上的关系的传递闭包。

a) {(a, c), (b, d), (c, a), (d, b), (e, d)}

b) {(b, c), (b, e), (c, e), (d, a), (e, b), (e, c)}

c) {(a, b), (a, c), (a, e), (b, a), (b, c), (c, a), (c, b), (d, a), (e, d)}

d) {(a, e), (b, a), (b, d), (c, d), (d, a), (d, c), (e, a), (e, b), (e, c), (e, e)}

27. 使用沃舍尔算法找出练习 25 中关系的传递闭包。
28. 使用沃舍尔算法找出练习 26 中关系的传递闭包。
29. 求出包含关系{(1, 2), (1, 4), (3, 3), (4, 1)}的最小的关系，使得它是

a) 自反的和传递的。　　b) 对称的和传递的。　　c) 自反的、对称的和传递的。

30. 完成引理 1 当 $a \neq b$ 的情况下的证明。
31. 已经设计出算法用 $O(n^{2.8})$ 次位运算来计算两个 $n \times n$ 的 0-1 矩阵的布尔积。假设可以使用这些算法，给

出用算法1和沃舍尔算法求 n 元素集合上关系的传递闭包所用位运算数的大 O 估计。

*32. 如果有向图的两个顶点间的最短路径存在，设计一个算法，利用路径中的内点的概念求这种最短路径的长度。

33. 修改算法1找出 n 元素集合上关系的传递闭包的自反闭包。

34. 修改沃舍尔算法找出 n 元素集合上关系的传递闭包的自反闭包。

35. 证明集合{0，1，2}上的关系 $R=\{(0,0),(0,1),(1,1),(2,2)\}$ 关于下述性质 **P** 的闭包不存在，如果 **P**

a)“不是自反的”。　　b)“有奇数个元素”。

7.5 等价关系

7.5.1 引言

在学期开始之前学院的学生登记注册。姓氏的第一个字母从A到G、从H到N和从O到Z的学生分别可以在上午8点到11点、上午11点到下午2点、下午2点到5点之间的任何时间注册。设 R 是关系，R 包含 (x, y)，当且仅当 x 和 y 是姓的第一个字母在同一组的学生。因此 x 和 y 可以在同一时间注册，当且仅当 (x, y) 属于 R。容易看出 R 是自反的、对称的和传递的。此外，根据他们的姓的第一个字母，R 把学生的集合划分成3类。为了解一个学生在什么时间注册，我们仅关心这个学生在3个类中的哪个类，我们并不需要识别这个学生。

如果4整除 $a-b$，整数 a 和 b 有模4同余的关系。后面我们将证明这个关系是自反的、对称的和传递的。不难看出 a 和 b 相关，当且仅当被4整除时，a 和 b 有相同的余数。这个关系将整数集划分成不同的4类。当我们仅关心一个整数被4整除的余数时，我们只需要知道它在哪个类而不必知道它的特定值。

这两个关系，R 和模4同余关系是等价关系，即是自反的、对称的和传递的。本节将证明这种关系把集合划分成由等价元素构成的不相交的类。当我们仅关心集合的一个元素是否在某个元素类，而不介意它的特殊身份时，就出现了等价关系。

7.5.2 等价关系

在这一节我们将研究具有一组特殊性质的关系，可以用这组性质为在某一方面类似的相关个体之间建立联系。

定义1 集合 A 上的关系叫做*等价关系*，如果它是自反的、对称的和传递的。

两个由等价关系联系起来的元素叫做*等价的元素*(因为等价关系是对称的，这个定义是有意义的)。由于等价关系是自反的，在一个等价关系中每个元素都与自身等价。此外，因为等价关系是传递的，如果 a 和 b 是等价的且 b 和 c 是等价的，那么 a 和 c 是等价的。

例1～例4说明了等价关系的概念。

例1 设 R 是英语字母串的集合上的关系并且使得 aRb 当且仅当 $l(a)=l(b)$，其中 $l(x)$ 是串 x 的长度。R 是等价关系吗?

解 因为 $l(a)=l(a)$，从而只要 a 是一个串，就有 aRa，故 R 是自反的。其次，假设 aRb，即 $l(a)=l(b)$。那么有 bRa，因为 $l(b)=l(a)$，因此 R 是对称的。最后，假设 aRb 和 bRc，那么有 $l(a)=l(b)$ 和 $l(b)=l(c)$。因此 $l(a)=l(c)$，即 aRc，从而 R 是传递的。由于 R 是自反的、对称的和传递的，R 是等价关系。■

例 2　设 R 是整数集上的关系，满足 aRb，当且仅当 $a=b$ 或 $a=-b$。在 7.1 节我们证明了 R 是自反的、对称的和传递的。因此 R 是等价关系。■

例 3　设 R 是实数集上的关系，满足 aRb，当且仅当 $a-b$ 是整数。R 是等价关系吗?

解　因为对所有的实数 a，$a-a=0$ 是整数，即对所有的实数有 aRa，因此 R 是自反的。现在假设 aRb，那么 $a-b$ 是整数，所以 $b-a$ 也是整数。因此有 bRa。R 是对称的。如果 aRb 和 bRc，那么 $a-b$ 和 $b-c$ 是整数，所以 $a-c=(a-b)+(b-c)$ 也是整数。因此 aRc。于是，R 是传递的。综上所述，R 是等价关系。■

最广泛使用的等价关系之一是模 m 同余关系，其中 m 是大于 1 的正整数。

例 4　**模 m 同余**。设 m 是大于 1 的正整数。证明关系

$$R=\{(a,b)\mid a\equiv b(\mathrm{mod}\ m)\}$$

是整数集上的等价关系。

解　回顾 2.4 节，$a\equiv b(\mathrm{mod}\ m)$，当且仅当 m 整除 $a-b$。注意 $a-a=0$ 被 m 整除，因为 $0=0\cdot m$。因此 $a\equiv a(\mathrm{mod}\ m)$，从而模 m 同余关系是自反的。现在假设 $a\equiv b(\mathrm{mod}\ m)$，那么 $a-b$ 被 m 整除，即 $a-b=km$，其中 k 是整数。从而 $b-a=(-k)m$，即 $b\equiv a(\mathrm{mod}\ m)$，因此模 m 同余关系是对称的。下面假设 $a\equiv b(\mathrm{mod}\ m)$ 和 $b\equiv c(\mathrm{mod}\ m)$，那么 m 整除 $a-b$ 和 $b-c$。因此，存在整数 k 和 l，使得 $a-b=km$ 和 $b-c=lm$。把这两个等式加起来得 $a-c=(a-b)+(b-c)=km+lm=(k+l)m$。于是，$a\equiv c(\mathrm{mod}\ m)$，从而模 m 同余关系是传递的。综上所述，模 m 同余关系是等价关系。■

7.5.3　等价类

设 A 是在你们学校所有的高中毕业生。考虑 A 上的关系 R，R 由所有的对 (x, y) 构成，其中 x 和 y 从同一高中毕业。给定学生 x，我们可以构造所有的关于 R 与 x 等价的学生的集合。这个集合由与 x 在同一高中毕业的所有学生构成。A 的这个子集叫做这个关系的一个等价类。

定义 2　设 R 是集合 A 上的等价关系。与 A 中的一个元素 a 有关系的所有元素的集合叫做 a 的等价类。A 的关于 R 的等价类记作 $[a]_R$。当只有一个关系被考虑时，我们将省去下标 R 并把这个等价类写作 $[a]$。

换句话说，如果 R 是集合 A 上的等价关系，元素 a 的等价类是

$$[a]_R=\{s\mid(a,s)\in R\}$$

如果 $b\in[a]_R$，b 叫做这个等价类的代表元。一个等价类的任何元素都可以作为这个类的代表元。也就是说，对作为这个类的代表元所选择的特定元素没有特殊要求。

例 5　对于例 2 的等价关系，一个整数的等价类是什么?

解　在这个等价关系中，一个整数等价于它自身和它的相反数。从而 $[a]=\{-a, a\}$。这个集合包含两个不同的整数，除非 $a=0$。例如，$[7]=\{-7, 7\}$，$[-5]=\{-5, 5\}$，$[0]=\{0\}$。■

例 6　对于模 4 同余关系，0 和 1 的等价类是什么?

解　0 的等价类包含使得 $a\equiv 0(\mathrm{mod}\ 4)$ 的所有整数 a。这个类中的整数是被 4 整除的那些整数。因此，对于这个关系 0 的等价类是

$$[0]=\{\cdots,-8,-4,0,4,8,\cdots\}$$

1 的等价类包含使得 $a\equiv 1(\mathrm{mod}\ 4)$ 的所有整数 a。这个类中的整数是当被 4 除时余数为 1 的那些整数。因此，对于这个关系 1 的等价类是

$$[1]=\{\cdots,-7,-3,1,5,9,\cdots\}$$

■

在例 6 找到了 0 和 1 关于模 4 同余的等价类。用任何正整数 m 代替 4 很容易把例 6 加以推

广。模 m 同余关系的等价类叫作**模 m 同余类**。整数 a 模 m 的同余类记作$[a]_m$，满足$[a]_m=\{\cdots, a-2m, a-m, a, a+m, a+2m, \cdots\}$。例如，从例 6 得出$[0]_4=\{\cdots, -8, -4, 0, 4, 8, \cdots\}$和$[1]_4=\{\cdots, -7, -3, 1, 5, 9, \cdots\}$。

7.5.4 等价类与划分

设 A 是你们学校恰好主修一个专业的学生的集合，R 是 A 上的关系，如果 x 和 y 是主修同一专业的学生，则(x, y)属于 R。那么正如读者可以验证的，R 是等价关系。我们可以看出 R 将 A 中的所有学生分成不相交的子集，其中每个子集包含某个特定专业的学生。例如，一个子集包含所有(只主修)计算机专业的学生，第二个子集包含所有历史专业的学生。而且这些子集是 R 的等价类。这个例子说明一个等价关系的等价类怎样把一个集合划分成不相交的非空子集。我们将在下面的讨论中把这些概念进一步精确化。

设 R 是集合 A 上的等价关系。下面的定理证明 A 的两个元素的等价类或是相等或是不相交。

定理 1 设 R 是集合 A 上的等价关系，下面的命题是等价的。

(i)aRb　　(ii)$[a]=[b]$　　(iii)$[a]\cap[b]\neq\varnothing$

证 首先证明(i)推出(ii)。假设 aRb，我们将通过$[a]\subseteq[b]$和$[b]\subseteq[a]$来证明$[a]=[b]$。假设 $c\in[a]$，那么 aRc。因为 aRb 和 R 的对称性，有 bRa。又由于 R 是传递的以及 bRa 和 aRc，就得到 bRc，因而有 $c\in[b]$。这就证明了$[a]\subseteq[b]$。类似地，可证明$[b]\subseteq[a]$，证明留给读者作为练习。

其次我们将证明(ii)推出(iii)。假设$[a]=[b]$。这就证明了$[a]\cap[b]\neq\varnothing$，因为$[a]$是非空的(由 R 的自反性 $a\in[a]$)。

下面证明(iii)推出(i)。假设$[a]\cap[b]\neq\varnothing$，那么存在元素 c 满足 $c\in[a]$和 $c\in[b]$。换句话说，aRc 和 bRc。由对称性有 cRb。再根据传递性，由 aRc 和 cRb 就有 aRb。

因为(i)推出(ii)，(ii)推出(iii)，(iii)推出(i)，三个命题(i)，(ii)和(iii)是等价的。 □

我们现在将显示一个等价关系怎样划分一个集合。设 R 是集合 A 上的等价关系，R 的所有等价类的并集就是 A 的全部，因为 A 的每个元素 a 都在它自己的等价类即$[a]_R$ 中。换句话说，

$$\bigcup_{a\in A}[a]_R=A$$

此外，由定理 1，这些等价类或是相等的或是不相交的，因此当$[a]_R\neq[b]_R$ 时，

$$[a]_R\cap[b]_R=\varnothing$$

这两个观察证明等价类构成 A 的划分，因为它们将 A 分成不相交的子集。更确切地说，集合 S 的划分是一族 S 的不相交的非空子集，且 S 就是它们的并。换句话说，一族子集 A_i，$i\in I$，(其中 I 是指标集)构成 S 的划分，当且仅当

$$A_i\neq\varnothing,\quad 对于 i\in I,$$
$$A_i\cap A_j=\varnothing, 当 i\neq j$$

和

$$\bigcup_{i\in I}A_i=S$$

(这里符号 $\bigcup_{i\in I}A_i$ 表示对于所有的 $i\in I$ 集合 A_i 的并集。)图 7-13 说明了集合的划分的概念。

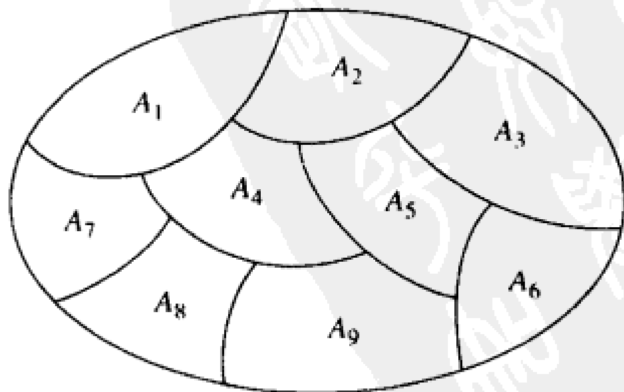


图 7-13 集合的划分

例 7 假设 $S=\{1, 2, 3, 4, 5, 6\}$，一族集合 $A_1=\{1, 2, 3\}$，$A_2=\{4, 5\}$和 $A_3=\{6\}$构成 S 的一个划分，因为这些集合是不相交的，且它们的并是 S。■

我们已经看到集合上等价关系的等价类构成这个集合的划分。这个划分中的子集就是这些等价类。反过来，可以用集合的每个划分来构成等价关系。两个元素关于这个关系是等价的，当且仅当它们在这个划分的同一子集中。

为看到这一点，假设$\{A_i \mid i\in I\}$是S的划分。设R是S上的由有序对(x, y)组成的等价关系，其中x和y属于这个划分的同一子集A_i。为证明R是等价关系，我们必须证明R是自反的、对称的和传递的。

对于每个$a\in S$有$(a, a)\in R$，因为a和它自己在同一子集中，因此R是自反的。如果$(a, b)\in R$，那么b和a在这个划分的同一子集中，因此有$(b, a)\in R$。从而R是对称的。如果$(a, b)\in R$和$(b, c)\in R$，那么a和b在这个划分的同一子集中，比如是X，b和c也在这个划分的同一子集中，比如是Y。因为划分的子集是不相交的，并且b属于X和Y，必有$X=Y$。因而a和c属于这个划分的同一子集，即$(a, c)\in R$。于是R是传递的。

这就证明了R是等价关系。R的等价类由S的子集构成，这些子集包含了S的有关系的元素，根据R的定义，它们就是划分的子集。定理2是对等价关系和划分之间建立的这种联系的总结。

定理2　设R是集合S上的等价关系。那么R的等价类构成S的划分。相反，给定集合S的划分$\{A_i \mid i\in I\}$，存在着等价关系R，它以集合$A_i(i\in I)$作为它的等价类。

例8说明了怎样从一个划分构造一个等价关系。

例8　$A_1=\{1, 2, 3\}$，$A_2=\{4, 5\}$，$A_3=\{6\}$是例7给出的集合$S=\{1, 2, 3, 4, 5, 6\}$的划分，列出这个划分所产生的等价关系R中的有序对。

解　划分中的子集是R的等价类。对$(a, b)\in R$，当且仅当a和b在划分的同一个子集中。由于$A_1=\{1, 2, 3\}$是一个等价类，因此，对$(1, 1)$，$(1, 2)$，$(1, 3)$，$(2, 1)$，$(2, 2)$，$(2, 3)$，$(3, 1)$，$(3, 2)$，$(3, 3)$属于R；由于$A_2=\{4, 5\}$是一个等价类，因此，对$(4, 4)$，$(4, 5)$，$(5, 4)$，$(5, 5)$也属于R；最后，由于$\{6\}$是一个等价类，对$(6, 6)$属于R。此外没有其他的对属于R。■

模m同余类对定理2提供了一个有用的说明。存在m个不同的模m同余类，对应于当一个整数除以m时可能得到的m个不同的余数。这m个同余类记作$[0]_m$，$[1]_m$，…，$[m-1]_m$。它们构成了整数集合的划分。

例9　模4同余产生的整数划分中的集合是什么？

解　存在4个同余类，对应于$[0]_4$，$[1]_4$，$[2]_4$和$[3]_4$，它们是集合

$$[0]_4=\{\cdots,-8,-4,0,4,8,\cdots\}$$
$$[1]_4=\{\cdots,-7,-3,1,5,9,\cdots\}$$
$$[2]_4=\{\cdots,-6,-2,2,6,10,\cdots\}$$
$$[3]_4=\{\cdots,-5,-1,3,7,11,\cdots\}$$

这些同余类是不相交的，并且每个整数恰好在它们之中的一个。换句话说，正如定理2所说，这些同余类构成了一个划分。■

练习

1. 下面是$\{0, 1, 2, 3\}$上的关系，其中哪些是等价关系？确定一个等价关系的性质，这些性质是其他关系所欠缺的。

   a) $\{(0, 0), (1, 1), (2, 2), (3, 3)\}$

   b) $\{(0, 0), (0, 2), (2, 0), (2, 2), (2, 3), (3, 2), (3, 3)\}$

c) $\{(0, 0), (1, 1), (1, 2), (2, 1), (2, 2), (3, 3)\}$

d) $\{(0, 0), (1, 1), (1, 3), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)\}$

e) $\{(0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (1, 2), (2, 0), (2, 2), (3, 3)\}$

2. 下面是所有人集合上的关系，其中哪些是等价关系？确定一个等价关系的性质，这些性质是其他关系所欠缺的。

   a) $\{(a, b) \mid a$ 与 b 有相同的年龄$\}$

   b) $\{(a, b) \mid a$ 与 b 有相同的父母$\}$

   c) $\{(a, b) \mid a$ 与 b 有一个相同的父亲或者一个相同的母亲$\}$

   d) $\{(a, b) \mid a$ 与 b 相识$\}$

   e) $\{(a, b) \mid a$ 与 b 说同一种语言$\}$

3. 下面是从 **Z** 到 **Z** 的所有函数集合上的关系，其中哪些是等价关系？确定一个等价关系的性质，这些性质是其他关系所欠缺的。

   a) $\{(f, g) \mid f(1)=g(1)\}$

   b) $\{(f, g) \mid f(0)=g(0)$ 或 $f(1)=g(1)\}$

   c) $\{(f, g) \mid$ 对所有的 $x\in \mathbf{Z}$，$f(x)-g(x)=1\}$

   d) $\{(f, g) \mid$ 对某个 $C\in \mathbf{Z}$，对所有的 $x\in \mathbf{Z}$，$f(x)-g(x)=C\}$

   e) $\{(f, g) \mid f(0)=g(1)$ 且 $f(1)=g(0)\}$

4. 定义 3 个在你们离散数学班中学生集合上的等价关系，要求与书中讨论的关系不同，确定关于这些等价关系的等价类。

5. 假设 A 是非空集合，f 是以 A 作为定义域的函数，设 R 是 A 上的关系，若 $f(x)=f(y)$，则 (x, y) 属于 R。

   a) 证明 R 是 A 上的等价关系。 b) R 的等价类是什么？

6. 假设 A 是非空集合，R 是 A 上的等价关系，证明存在以 A 作为定义域的函数 f，使得 $(x, y)\in R$，当且仅当 $f(x)=f(y)$。

7. 设 R 是长度至少为 3 的所有二进位串的集合上的关系，R 由对 (x, y) 构成，其中 x 和 y 在它们的前 3 位相同。证明 R 是等价关系。

8. 设 R 是长度至少为 3 的所有二进位串的集合上的关系，R 由对 (x, y) 构成，其中 x 和 y 除了在它们的前 3 位有可能不同之外其他位都相同。证明 R 是等价关系。

9. 证明命题等价是在所有复合命题集合上的等价关系。

10. 设 R 是正整数的有序对集合上的关系，$((a, b), (c, d))\in R$ 当且仅当 $ad=bc$。证明 R 是等价关系。

11. (需要微积分)

    a) 设 R 是从 **R** 到 **R** 的所有可微分函数的集合上的关系，R 由所有的对 (f, g) 构成，其中对所有实数 x，$f'(x)=g'(x)$。证明 R 是等价关系。

    b) 什么函数与函数 $f(x)=x^2$ 在同一个等价类中？

12. (需要微积分)

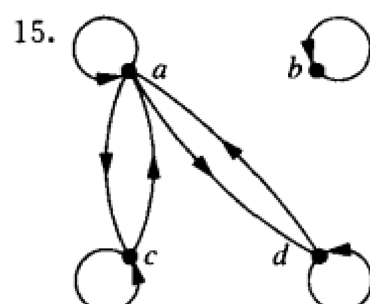
    a) 设 n 是正整数，R 是实系数多项式集合上的关系，R 由所有的对 (f, g) 构成，其中 $f^{(n)}(x)=g^{(n)}(x)$ [这里的 $f^{(n)}(x)$ 是 $f(x)$ 的 n 阶导数]，证明 R 是等价关系。

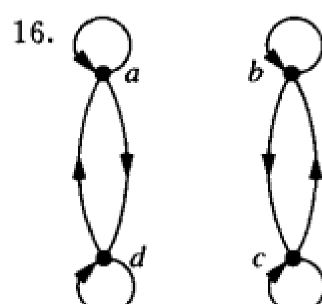
    b) 什么函数与函数 $f(x)=x^4$ 在同一个等价类中，其中 $n=3$？

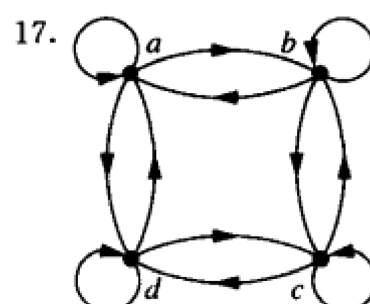
13. 设 R 是所有 URL(或 Web 地址)集合上的关系，xRy 当且仅当在 x 的 Web 页与在 y 的 Web 页相同，证明 R 是等价关系。

14. 设 R 是已经访问过某个特定 Web 页的所有人的集合上的关系，xRy 当且仅当网页浏览者 x 和网页浏览者 y 从这个网页开始按照同样的一组链接进行访问(从一个 Web 页跳转到另一个 Web 页直到他们停止使用 Web)。证明 R 是等价关系。

在练习 15～17 中确定具有所示有向图的关系是否为等价关系。

15. 

16. 

17. 

18. 确定由下面的 0-1 矩阵表示的关系是否为等价关系。

a) $\begin{bmatrix}1&1&1\\0&1&1\\1&1&1\end{bmatrix}$ b) $\begin{bmatrix}1&0&1&0\\0&1&0&1\\1&0&1&0\\0&1&0&1\end{bmatrix}$ c) $\begin{bmatrix}1&1&1&0\\1&1&1&0\\1&1&1&0\\0&0&0&1\end{bmatrix}$

19. 设 R 是所有二进位串的集合上的关系，sRt 当且仅当 s 和 t 包含相同个数的 1，证明 R 是等价关系。

20. 练习 1 中的等价关系的等价类是什么？

21. 练习 2 中的等价关系的等价类是什么？

22. 练习 3 中的等价关系的等价类是什么？

23. 对于练习 19 中的等价关系，二进位串 011 的等价类是什么？

24. 对于练习 7 中的等价关系，下述二进位串的等价类是什么？

a)010 b)1011 c)11111 d)01010101

25. 对于练习 8 中的等价关系，描述练习 24 中的二进制串的等价类。

26. 当 m 是下面的整数时，$[4]_m$ 的同余类是什么？

a)2 b)3 c)6 d)8

27. 描述每一个模 6 同余类。

28. a)对于练习 10 中的等价关系(1，2)的等价类是什么？

b)对于练习 10 中的等价关系 R 解释等价类的含义。

29. 下面哪些子集族是{1，2，3，4，5，6}的划分？

a){1，2}，{2，3，4}，{4，5，6} b){1}，{2，3，6}，{4}，{5}

c){2，4，6}，{1，3，5} d){1，4，5}，{2，6}

30. 下面哪些子集族是{−3，−2，−1，0，1，2，3}的划分？

a){−3，−1，1，3}，{−2，0，2} b){−3，−2，−1，0}，{0，1，2，3}

c){−3，3}，{−2，2}，{−1，1}，{0} d){−3，−2，2，3}，{−1，1}

31. 下面哪些子集族是 8 位二进位串的集合上的划分？

a)以 1 开始的二进位串的集合，以 00 开始的二进位串的集合，以 01 开始的二进位串的集合。

b)包含串 00 的二进位串的集合，包含串 01 的二进位串的集合，包含串 10 的二进位串的集合，包含串 11 的二进位串的集合。

c)以 00 结尾的二进位串的集合，以 01 结尾的二进位串的集合，以 10 结尾的二进位串的集合，以 11 结尾的二进位串的集合。

d)以 111 结尾的二进位串的集合，以 011 结尾的二进位串的集合，以 00 结尾的二进位串的集合。

e)含 $3k$ 个 1 的二进位串的集合，其中 k 为非负整数；含 $3k+1$ 个 1 的二进位串的集合，其中 k 为非负整数；含 $3k+2$ 个 1 的二进位串的集合，其中 k 是正整数。

32. 下面哪些子集族是整数集合的划分？

a)偶数集合与奇数集合。

b)正整数集合与负整数集合。

c)被 3 整除的整数集合，当被 3 除时余数为 1 的整数集合，当被 3 除时余数为 2 的整数集合。

d)小于 −100 的整数集合，绝对值不超过 100 的整数集合，大于 100 的整数集合。

e)不能被3整除的整数集合，偶数集合，当被6除时余数为3的整数集合。

33. 下面哪些是整数的有序对的集合 $\mathbf{Z}\times\mathbf{Z}$ 的划分？

a) x 或 y 是奇数的 (x, y) 对的集合；x 是偶数的 (x, y) 对的集合；y 是偶数的 (x, y) 对的集合。

b) x 和 y 都是奇数的 (x, y) 对的集合；x 和 y 只有一个是奇数的 (x, y) 对的集合；x 和 y 都是偶数的 (x, y) 对的集合。

c) x 是正数的 (x, y) 对的集合；y 是正数的 (x, y) 对的集合；x 和 y 都是负数的 (x, y) 对的集合。

d) x 和 y 都被3整除的 (x, y) 对的集合；x 被3整除且 y 不被3整除的 (x, y) 对的集合；x 不被3整除且 y 被3整除的 (x, y) 对的集合；x 和 y 都不被3整除的 (x, y) 对的集合。

e) $x>0$ 且 $y>0$ 的 (x, y) 对的集合；$x>0$ 且 $y\leqslant 0$ 的 (x, y) 对的集合；$x\leqslant 0$ 且 $y>0$ 的 (x, y) 对的集合；$x\leqslant 0$ 且 $y\leqslant 0$ 的 (x, y) 对的集合。

f) $x\neq 0$ 且 $y\neq 0$ 的 (x, y) 对的集合；$x=0$ 且 $y\neq 0$ 的 (x, y) 对的集合；$x\neq 0$ 且 $y=0$ 的 (x, y) 对的集合。

34. 下面哪些是实数集合的划分？

a)负实数集合，$\{0\}$，正实数集合[⊖]

b)无理数集合，有理数集合

c)区间 $[k, k+1]$，$k=\cdots, -2, -1, 0, 1, 2, \cdots$

d)区间 $(k, k+1)$，$k=\cdots, -2, -1, 0, 1, 2, \cdots$

e)区间 $(k, k+1]$，$k=\cdots, -2, -1, 0, 1, 2, \cdots$

f)集合 $\{x+n \mid n\in\mathbf{Z}\}$，对一切 $x\in[0, 1)$

35. 列出由 $\{0, 1, 2, 3, 4, 5\}$ 的划分产生的等价关系中的有序对。

a) $\{0\}$，$\{1, 2\}$，$\{3, 4, 5\}$　　b) $\{0, 1\}$，$\{2, 3\}$，$\{4, 5\}$

c) $\{0, 1, 2\}$，$\{3, 4, 5\}$　　d) $\{0\}$，$\{1\}$，$\{2\}$，$\{3\}$，$\{4\}$，$\{5\}$

36. 列出由 $\{a, b, c, d, e, f, g\}$ 的划分产生的等价关系中的有序对。

a) $\{a, b\}$，$\{c, d\}$，$\{e, f, g\}$　　b) $\{a\}$，$\{b\}$，$\{c, d\}$，$\{e, f\}$，$\{g\}$

c) $\{a, b, c, d\}$，$\{e, f, g\}$　　d) $\{a, c, e, g\}$，$\{b, d\}$，$\{f\}$

如果在划分 P_1 中的每个集合都是划分 P_2 中每个集合的子集，则 P_1 叫做 P_2 的**加细**。

37. 证明由模6同余类构成的划分是模3同余类构成的划分的加细。

38. 证明对于住在美国的人的集合，由住在同一郡或教区的人的子集构成的划分是住在同一州的人的子集所构成划分的加细。

39. 证明对于16位二进位串的集合，最后8位相同的二进位串的等价类所构成的划分是由最后4位相同的二进位串的等价类所构成的划分的加细。

40. 假设 R_1 和 R_2 是集合 A 上的等价关系，P_1 和 P_2 分别是对应于 R_1 和 R_2 的划分。证明 $R_1\subseteq R_2$，当且仅当 P_1 是 P_2 的加细。

41. 求出在集合 $\{a, b, c, d, e\}$ 上包含关系 $\{(a, b), (a, c), (d, e)\}$ 的最小的等价关系。

42. 假设 R_1 和 R_2 是集合 S 上的等价关系。确定下面 R_1 与 R_2 的每个组合是否一定为等价关系。

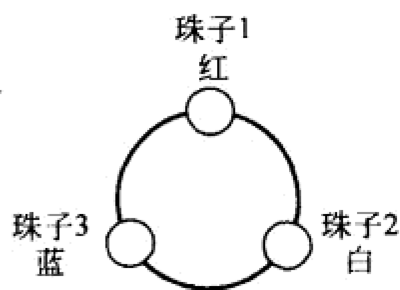
a) $R_1\cup R_2$　　b) $R_1\cap R_2$　　c) $R_1\oplus R_2$

43. 考虑例3的等价关系，即 $R=\{(x, y) \mid x-y$ 是整数$\}$。

a)关于这个等价关系的1的等价类是什么？

b)关于这个等价关系的1/2的等价类是什么？

*44. 如插图所示，在具有3颗珠子的手镯上每颗珠子可以是红的、白的或蓝的。如下定义手镯之间的等价关系 R：设 B_1 和 B_2 是手镯，(B_1, B_2) 属于



⊖ 原文误为“负实数，{0}，正实数”，已改正。——译者注

R，当且仅当 B_2 可以由旋转 B_1 得到，或先旋转 B_1 然后再翻转 B_1 得到。

a)证明 R 是等价关系。　　b)R 的等价类是什么？

*45. 设 R 是 2×2 棋盘的所有涂色集合上的关系，其中 4 个方格中的每一个可以被涂成红色或蓝色。C_1 和 C_2 是被这样涂色的 2×2 棋盘，(C_1, C_2)属于 R，当且仅当 C_2 可以由旋转 C_1，或旋转 C_1 然后再翻转 C_1 得到。

a)证明 R 是等价关系。　　b)R 的等价类是什么？

46. a)设 R 是从 $\mathbf{Z}^+$ 到 $\mathbf{Z}^+$ 的函数的集合上的关系，且$(f, g)\in R$，当且仅当 f 是 $\Theta(g)$(参见 2.2 节)。证明 R 是等价关系。

b)对于(a)的等价关系，描述包含 $f(n)=n^2$ 的等价类。

47. 通过列举确定 3 元集上的不同的等价关系个数。

48. 通过列举确定 4 元集上的不同的等价关系个数。

*49. 当我们构造一个关系的自反闭包的对称闭包的传递闭包时，一定得到一个等价关系吗？

*50. 当我们构造一个关系的传递闭包的自反闭包的对称闭包时，一定得到一个等价关系吗？

51. 假设我们使用定理 2 从一个等价关系 R 构造一个划分 P。如果再次使用定理 2 从 P 构造一个等价关系，那么结果的等价关系 R'是什么？

52. 假设我们使用定理 2 从一个划分 P 构造一个等价关系 R。如果再次使用定理 2 从 R 构造一个划分，那么结果的划分 P'是什么？

53. 设计一个算法找出包含一个给定关系的最小的等价关系。

*54. 设 $p(n)$表示 n 元集上的不同的等价关系个数(由定理 2，也是 n 元素集合的划分个数)。证明 $p(n)$满足递推关系 $p(n)=\sum_{j=0}^{n-1} C(n-1, j)p(n-j-1)$和初始条件 $p(0)=1$。[注：数 $p(n)$叫做贝尔数，根据美国数学家 E. T. 贝尔命名。]

55. 用练习 54 求 n 元素集合上的不同的等价关系个数，其中 n 是不超过 10 的正整数。

7.6 偏序

7.6.1 引言

我们常常用关系对集合的某些元素或全体元素排序。例如，使用包含着字对(x, y)的关系对字排序，其中 x 按照字典顺序排在 y 的前面。使用包含着对(x, y)的关系安排课题，其中 x 和 y 是课题中的任务并且 x 必须在 y 开始之前完成。使用包含着对(x, y)的关系排序整数集合，其中 x 小于 y。当我们把所有形如(x, x)的对加到这些关系中时，就得到了一个自反、反对称和传递的关系。这些都是刻画关系特征的性质，而这些关系使用它们相对的大小来排序集合的元素。

定义 1　集合 S 上的关系R，如果它是自反的、反对称的和传递的，就称为偏序。集合 S 与偏序 R 一起叫作偏序集，记作(S, R)。

例 1　证明"大于或等于"关系($\geqslant$)是整数集合上的偏序。

解　因为对所有整数 a 有 $a\geqslant a$，$\geqslant$是自反的。如果 $a\geqslant b$ 且 $b\geqslant a$，那么 $a=b$，因此$\geqslant$是反对称的。最后，因为 $a\geqslant b$ 和 $b\geqslant c$ 推出 $a\geqslant c$，所以$\geqslant$是传递的。从而$\geqslant$是整数集合上的偏序且$(\mathbf{Z}, \geqslant)$是偏序集。■

例 2　整除关系"|"是正整数集合上的偏序，因为正如 7.1 节所述，它是自反的、反对称的和传递的。我们看到$(\mathbf{Z}^+, |)$是偏序集($\mathbf{Z}^+$表示正整数集合)。■

例 3　证明包含关系$\subseteq$是集合 S 的幂集上的偏序。

解　因为只要 A 是 S 的子集，就有 $A\subseteq A$，$\subseteq$是自反的。因为 $A\subseteq B$ 和 $B\subseteq A$ 推出 $A=B$，

因此它是反对称的。最后，因为 $A\subseteq B$ 和 $B\subseteq C$ 推出 $A\subseteq C$，$\subseteq$ 是传递的。因此，$\subseteq$ 是 $P(S)$ 上的偏序，且 $(P(S), \subseteq)$ 是偏序集。■

在一个偏序集中记号 $a\preccurlyeq b$ 表示 $(a, b)\in R$。使用这个记号是由于"小于或等于"关系是偏序关系的范例。(注意符号 $\preccurlyeq$ 用来表示任意偏序集的关系，并不仅仅是"小于或等于"关系。)记号 $a\prec b$ 表示 $a\preccurlyeq b$，但 $a\neq b$。如果 $a\prec b$，我们说"a 小于 b"或"b 大于 a"。

当 a 与 b 是偏序集 $(S, \preccurlyeq)$ 的元素时，不一定有 $a\preccurlyeq b$ 或 $b\preccurlyeq a$。例如，在 $(P(\mathbf{Z}), \subseteq)$ 中，$\{1, 2\}$ 与 $\{1, 3\}$ 没有关系，反之亦然，因为没有一个集合被另一个集合包含。类似地，在 $(\mathbf{Z}, |)$ 中，2 与 3 没关系，3 也与 2 没关系，因为 $2\nmid 3$ 且 $3\nmid 2$. 由此得到定义 2。

定义 2 偏序集 $(S, \preccurlyeq)$ 的元素 a 和 b 叫作*可比的*，如果 $a\preccurlyeq b$ 或 $b\preccurlyeq a$。当 a 和 b 是 S 的元素并且既没有 $a\preccurlyeq b$，也没有 $b\preccurlyeq a$，则称 a 与 b 是*不可比的*。

例 4 在偏序集 $(\mathbf{Z}^+, |)$ 中，整数 3 和 9 是可比的吗？5 和 7 是可比的吗？

解 整数 3 和 9 是可比的，因为 $3\mid 9$。整数 5 和 7 是不可比的，因为 $5\nmid 7$ 且 $7\nmid 5$。■

用形容词"部分的(偏的)"描述偏序是由于一些对元素可能是不可比的。当集合中的每对元素都可比时，这个关系叫作全序。

定义 3 如果 $(S, \preccurlyeq)$ 是偏序集，且 S 的每对元素都是可比的，则 S 叫作*全序集*或*线序集*，且 $\preccurlyeq$ 叫作*全序*或*线序*。一个全序集也叫作*链*。

例 5 偏序集 $(\mathbf{Z}, \leqslant)$ 是全序集，因为只要 a 和 b 是整数，就有 $a\leqslant b$ 或 $b\leqslant a$。■

例 6 偏序集 $(\mathbf{Z}^+, |)$ 不是全序集，因为它包含着不可比的元素，例如 5 和 7。■

在第 3 章我们注意到 $(\mathbf{Z}^+, \leqslant)$ 是良序的，其中 $\leqslant$ 是通常的"小于或等于"关系。我们现在定义良序集。

定义 4 对于偏序集 $(S, \preccurlyeq)$，如果 $\preccurlyeq$ 是全序，并且 S 的每个非空子集都有一个最小元素，就称它为*良序集*。

例 7 正整数的有序对的集合 $\mathbf{Z}^+\times\mathbf{Z}^+$ 与 $\preccurlyeq$ 构成良序集。对于 (a_1, a_2) 和 (b_1, b_2)，如果 $a_1<b_1$，或如果 $a_1=b_1$ 且 $a_2\leqslant b_2$(字典顺序)，则 $(a_1, a_2)\preccurlyeq(b_1, b_2)$。有关的验证留作节后的练习。集合 $\mathbf{Z}$ 与通常的 $\leqslant$ 序不是良序的，因为负整数集合是 $\mathbf{Z}$ 的子集，但没有最小元素。■

在 3.4 节我们显示了怎样使用良序归纳原理证明关于一个良序集的结果。现在我们叙述并证明这个证明技术是有效的。

定理 1 *良序归纳原理*。设 S 是一个良序集，如果下述条件成立：

基础步：$P(x_0)$ 对 S 的最小元素为真，且

归纳步：对一切 $y\in S$，如果 $P(x)$ 对所有的 $x\prec y$ 为真，则 $P(y)$ 为真。

那么 $P(x)$ 对所有的 $x\in S$ 为真。

证 假若 $P(x)$ 不对所有的 $x\in S$ 为真。那么存在一个元素 $y\in S$ 使得 $P(y)$ 为假。于是集合 $A=\{x\in S\mid P(x)\text{为假}\}$ 是非空的。因为 S 是良序的，集合 A 有最小元素 a。易知 $a\neq x_0$，因为从基础步知道 $P(x_0)$ 为真。根据 a 是选自 A 的最小元素，我们知道对所有的 $x\in S(x\prec a)$ 都有 $P(x)$ 为真。由归纳步这就可以推出 $P(a)$ 为真。这个矛盾就证明了 $P(x)$ 必须对所有 $x\in S$ 为真。□

良序归纳原理对证明关于良序集的结果是一种通用的技术。即使可以使用关于正整数集合的数学归纳法证明一个定理时，使用良序归纳原理甚至可能更简单。正如我们在 3.4 节例 15 中所看到的，在那里我们证明了一个关于良序集 $(\mathbf{N}\times\mathbf{N}, \preccurlyeq)$ 的结果，其中 $\preccurlyeq$ 是 $\mathbf{N}\times\mathbf{N}$ 上的字典顺序。

7.6.2 字典顺序

字典中的字按照字母顺序或字典顺序排列，字典顺序是以字母表中的字母顺序为基础的。

这是从一个集合上的偏序构造一个集合上的串的序的特殊情况。我们将说明这种构造在任一个偏序集上是怎样做的。

首先，我们将说明怎样在两个偏序集$(A_1, \preccurlyeq_1)$和$(A_2, \preccurlyeq_2)$的笛卡儿积上构造一个偏序。在$A_1 \times A_2$上的字典顺序定义如下：如果第一个对的第一个元素（在A_1中）小于第二个对的第一个元素，或者第一个元素相等，但是第一个对的第二个元素（在A_2中）小于第二个对的第二个元素，那么第一个对小于第二个对。换句话说，(a_1, a_2)小于(b_1, b_2)，即

$$(a_1, a_2) \prec (b_1, b_2),$$

或者$a_1 \prec_1 b_1$，或者$a_1 = b_1$且$a_2 \prec_2 b_2$。

把相等加到$A_1 \times A_2$上的序$\prec$就得到偏序$\preccurlyeq$。这个验证留作练习。

例 8　确定在偏序集$(\mathbf{Z} \times \mathbf{Z}, \preccurlyeq)$中是否有$(3, 5) \prec (4, 8)$，$(3, 8) \prec (4, 5)$和$(4, 9) \prec (4, 11)$？这里的$\preccurlyeq$是从$\mathbf{Z}$上通常的$\leqslant$关系构造的字典顺序。

解　因为$3 < 4$，故而$(3, 5) \prec (4, 8)$且$(3, 8) \prec (4, 5)$。因为$(4, 9)$与$(4, 11)$的第一元素相同，但是$9 < 11$，我们有$(4, 9) \prec (4, 11)$。■

图 7-14 明显地显示了$\mathbf{Z}^+ \times \mathbf{Z}^+$中比$(3, 4)$小的有序对的集合。

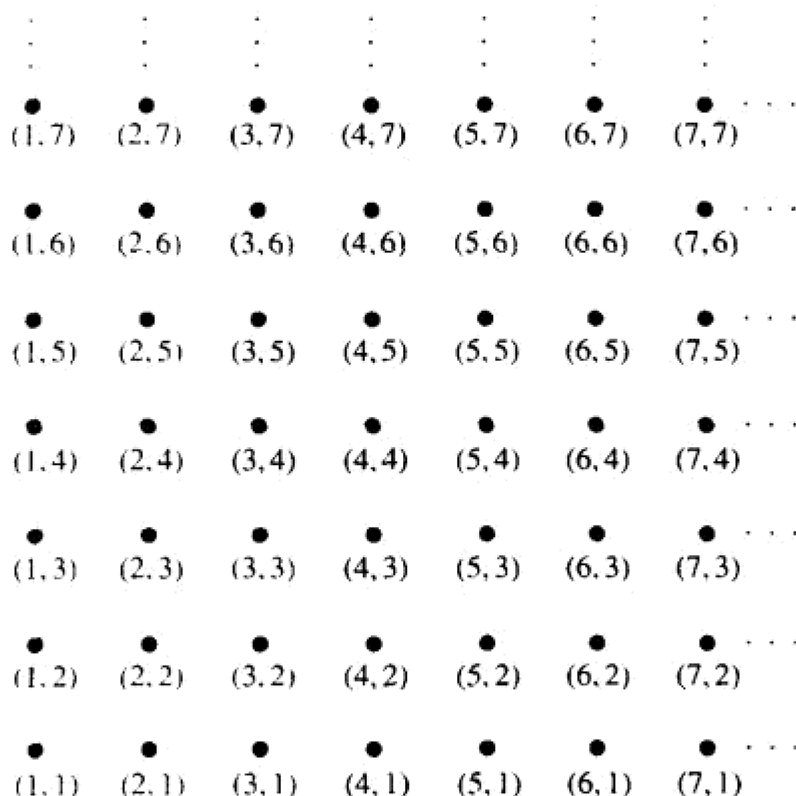


图 7-14　按照字典顺序比(3，4)小的有序对

可以在n个偏序集$(A_1, \preccurlyeq_1)$，$(A_2, \preccurlyeq_2)$，…，$(A_n, \preccurlyeq_n)$的笛卡儿积上定义字典顺序。如下定义$A_1 \times A_2 \times \cdots \times A_n$上的偏序$\preccurlyeq$：如果$a_1 \prec b_1$，或者存在整数$i > 0$使得$a_1 = b_1$，…，$a_i = b_i$，且$a_{i+1} \prec_{i+1} b_{i+1}$，那么

$$(a_1, a_2, \cdots, a_n) \prec (b_1, b_2, \cdots, b_n)$$

换句话说，如果在两个n元组不同元素出现的第一位置上第一个n元组的元素小于第二个n元组的元素，那么第一个n元组小于第二个n元组。

例 9　注意$(1, 2, 3, 5) \prec (1, 2, 4, 3)$，因为这些 4 元组的前两位相同，但是第一个 4 元组的第三位 3 小于第二个 4 元组的第三位 4（这里的 4 元组上的字典顺序是整数集合上的通常的“小于或等于”关系导出的字典顺序）。■

我们现在可以定义串上的字典顺序。考虑偏序集S上的串$a_1a_2\cdots a_m$和$b_1b_2\cdots b_n$，假定这些串不相等。设t是m，n中较小的数。定义字典顺序如下：$a_1a_2\cdots a_m$小于$b_1b_2\cdots b_n$，当且仅当

$$(a_1, a_2, \cdots, a_t) \prec (b_1, b_2, \cdots, b_t) \quad \text{或者}$$
$$(a_1, a_2, \cdots, a_t) = (b_1, b_2, \cdots, b_t) \text{ 并且 } m < n$$

其中这个不等式中的$\prec$表示S^t中的字典顺序。换句话说，为确定两个不同串的序，较长的串被切到较短的串的长度t，即$t=\min(m, n)$。然后使用S^t上的字典顺序比较每个串的前t位组成的t元组。如果对应于第一个串的t元组小于第二个串的t元组，或者这两个t元组相等，但是第二个串更长，那么第一个串小于第二个串。这是偏序的验证留给读者作为练习。

例 10 考虑小写英语字母串构成的集合。使用在字母表中的字母序，可以构造在串的集合上的字典顺序。如果两个串第一次出现不同字母时，第一个串的字母先于第二个串的字母，或者如果第一个串和第二个串在所有的位都相同，但是第二个串有更多的字母，那么，第一个串小于第二个串。这种排序和字典使用的排序相同。例如，

$$discreet \prec discrete$$

因为这两个串在第7位首次出现不同字母，并且$e \prec t$。

$$discreet \prec discreetness$$

因为这两个串前8个字母相同，但是第二个串更长。此外，

$$discrete \prec discretion$$

因为

$$discrete \prec discreti$$

■

7.6.3 哈塞图

在有穷偏序集的有向图中有许多边可以不必显示出来，因为它们是必须存在的。例如，考虑在集合$\{1, 2, 3, 4\}$上的偏序$\{(a, b) \mid a \leq b\}$的有向图，参见图7-15a。因为这个关系是偏序，它是自反的并且有向图在所有的顶点都有环。因此，我们不必显示这些环，因为它们是必须出现的。在图7-15b中没有显示这些环。由于一个偏序是传递的，我们不必显示那些由于传递性而必须出现的边。例如，在图7-15c中没有显示边(1, 3)，(1, 4)和(2, 4)，因为它们是必须出现的。如果假设所有边的方向是向上的(如图中所示)，我们不必显示边的方向；图7-15c没有显示方向。

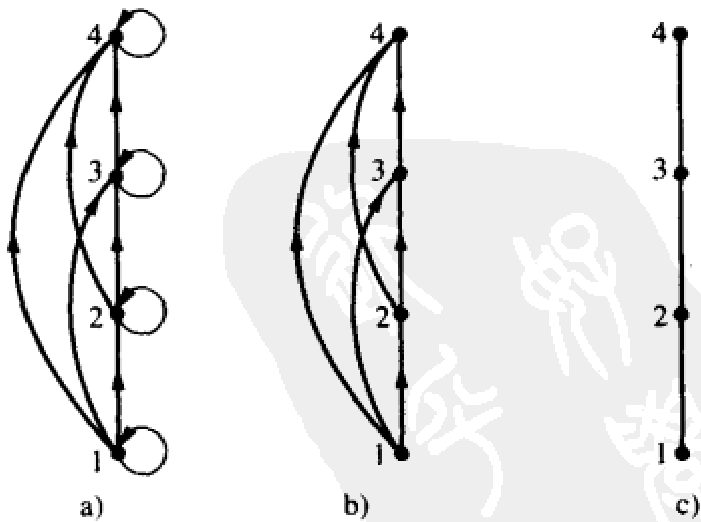


图7-15 构造关于$(\{1, 2, 3, 4\}, \leq)$的哈塞图

一般说来，我们可以使用下面的过程表示一个有穷集上的偏序。从这个关系的有向图开始。由于偏序是自反的，每个顶点有一个环，移走这些环。移走所有由于传递性出现的边，因为偏序是传递的，这些边是必须出现的。例如，如果(a, b)和(b, c)在偏序中，移走边(a, c)，因为它必须要出现。此外，如果(c, d)也在偏序中，移走边(a, d)，因为它也是必须出现的。最后，排列每条边使得它的起点在终点的下面(正如在纸上所画的)。移走所有有向边上的箭头，因为所有的边向上指向它们的终点。(只有和偏序集的覆盖关系中的对所对应的边留下来。见节末练习22前面的说明。)

这些步骤是有着明确定义的，并且对于一个有穷偏序集只有有限步需要执行。当所有的步骤执行以后，就得到一个包含着足够的表示偏序信息的图。这个图叫作哈塞图，它是用20世纪

德国数学家赫尔姆·哈塞[⊖]的名字命名的。

例 11　画出表示{1, 2, 3, 4, 6, 8, 12}上的偏序{(a, b) | a 整除 b}的哈塞图。

解　由这个偏序的有向图开始，如图 7-16a 所示。移走所有的环，如图 7-16b 所示。然后删除所有由传递性导出的边。这些边是(1, 4)，(1, 6)，(1, 8)，(1, 12)，(2, 8)，(2, 12)和(3, 12)。排列所有的边使得方向向上，并且删除所有的箭头得到哈塞图。结果如图 7-16c 所示。■

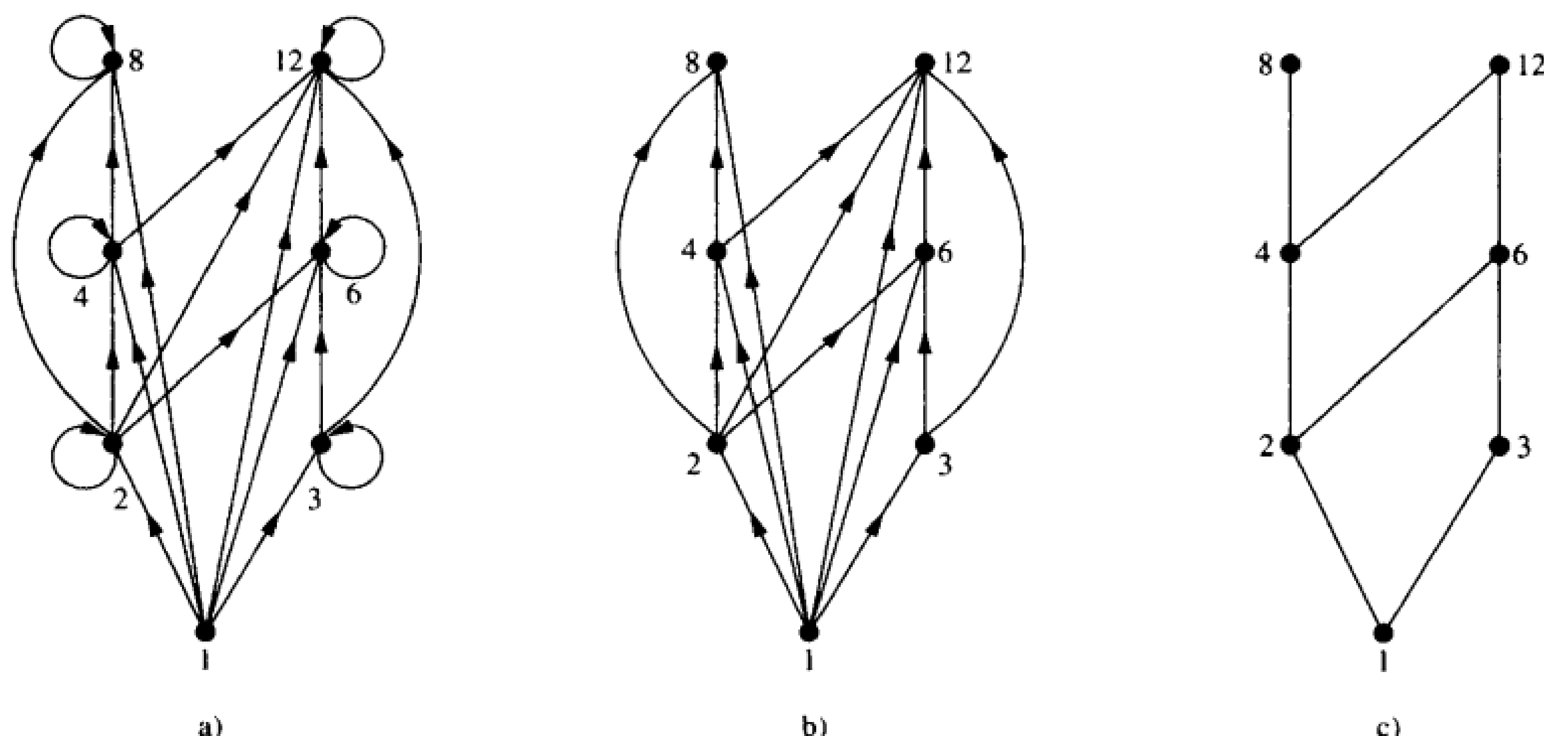


图 7-16　构造({1, 2, 3, 4, 6, 8, 12}, |)上的哈塞图

例 12　画出幂集 $P(S)$上的偏序{(A, B) | $A\subseteq B$}的哈塞图，其中 $S=\{a, b, c\}$。

解　关于这个偏序的哈塞图是由相关的有向图得到的，先删除所有的环和所有由传递性产生的边，即$(\varnothing, \{a, b\})$，$(\varnothing, \{a, c\})$，$(\varnothing, \{b, c\})$，$(\varnothing, \{a, b, c\})$，$(\{a\}, \{a, b, c\})$，$(\{b\}, \{a, b, c\})$和$(\{c\}, \{a, b, c\})$。最后，使所有的边方向向上并删除箭头。得到的哈塞图如图 7-17 所示。■

7.6.4　极大元素与极小元素

具有极值性质的偏序集元素有许多重要应用。偏序集的一个元素叫作极大的，当它不小于这个偏序集的任何其他元素。即 a 在偏序集$(S, \preccurlyeq)$中是*极大的*，当不存在 $b\in S$ 使得 $a\prec b$。类似地，偏序集的一个元素叫作极小的，如果它不大于这个偏序集的任何其他元素。即 a 在偏序集$(S, \preccurlyeq)$中是*极小的*，如果不存在 $b\in S$ 使得 $b\prec a$。使用哈塞图很容易识别极大元素与极小元素。它们是图中的"顶"元素与"底"元素。

例 13　偏序集({2, 4, 5, 10, 12, 20, 25}, |)的哪些元素是极大的，哪些是极小的？

⊖ **赫尔姆·哈塞(Helmut Hasse, 1898—1979)**　哈塞生于卡斯尔，高中毕业以后在德国海军服役。他于 1918 年进入哥廷根大学学习，并于 1920 年到马尔堡大学在数论专家科特·亨泽尔的指导下学习。在这段时间，哈塞对代数数论做出了基础性的贡献。他接替了亨泽尔在马尔堡大学的工作，1934 年成为著名的哥廷根数学研究所的所长，并在 1950 年受聘于汉堡大学。哈塞为著名的德国数学期刊《Crelle 学报》担任了 50 年的编辑工作。当纳粹强迫亨泽尔辞职时，他在 1936 年承担了主编的工作。在第二次世界大战时期，哈塞为德国海军从事应用数学研究。他的讲演表述清晰，风格独特。他把一生的精力都倾注于数论和他的学生。(哈塞由于与纳粹党的联系而受到非议。调查证明他是强烈的日耳曼民族主义者，但不是狂热的纳粹分子。)

解 图 7-18 关于这个偏序集的哈塞图显示了极大元素是 12，20 和 25，极小元素是 2 和 5。通过这个例子可以看出，一个偏序集可以有多于一个的极大元素和多于一个的极小元素。■

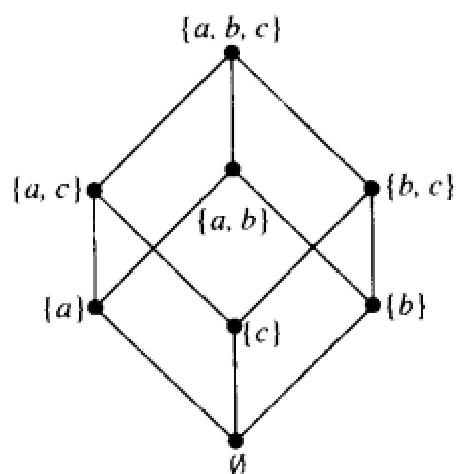


图 7-17 $(P(\{a, b, c\}), \subseteq)$的哈塞图

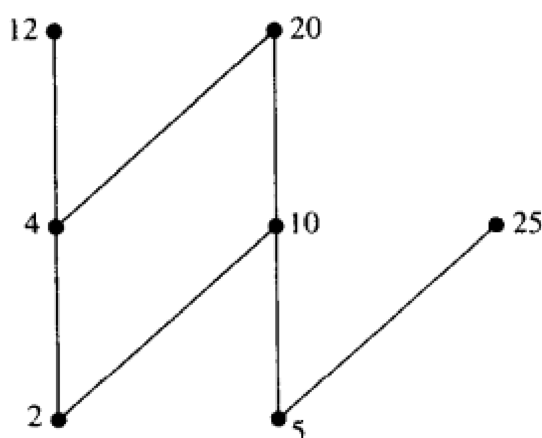


图 7-18 偏序集的哈塞图

有时在偏序集中存在一个元素大于每个其他的元素。这样的元素叫做最大元素。即 a 是偏序集$(S, \preccurlyeq)$的**最大元素**，当对所有的 $b \in S$ 有 $b \preccurlyeq a$。当最大元素存在时它是唯一的[见节末练习 34(a)]。类似地，一个元素叫做最小元素，当它小于偏序集的所有其他元素。即 a 是偏序集$(S, \preccurlyeq)$的**最小元素**，如果对所有的 $b \in S$ 有 $a \preccurlyeq b$。当最小元素存在时它也是唯一的[见节末练习 34(b)]。

例 14 确定图 7-19 的每个哈塞图表示的偏序集是否有最大元素和最小元素。

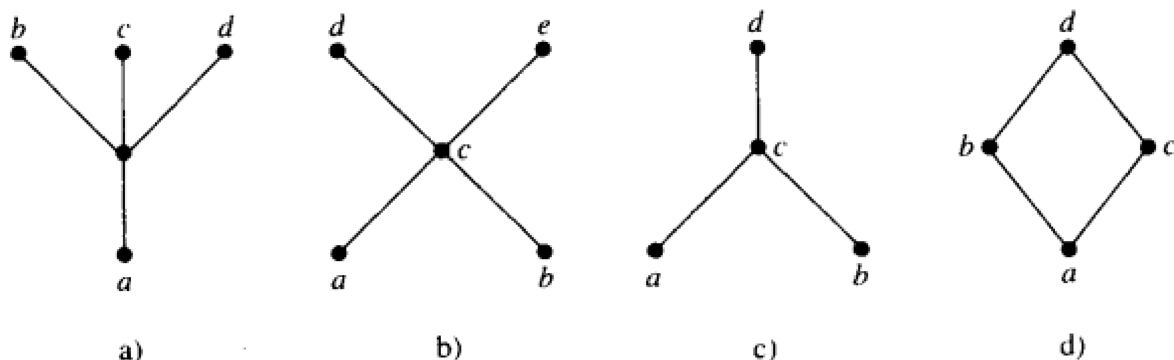


图 7-19 四个偏序集的哈塞图

解 哈塞图图 7-19a 的偏序集的最小元素是 a。这个偏序集没有最大元素。哈塞图图 7-19b 的偏序集既没有最小元素也没有最大元素。哈塞图图 7-19c 的偏序集没有最小元素，它的最大元素是 d。哈塞图图 7-19d 的偏序集有最小元素 a 和最大元素 d。■

例 15 设 S 是集合。确定偏序集$(P(S), \subseteq)$中是否存在最大元素与最小元素。

解 最小元素是空集，因为对于 S 的任何子集 T，有 $\varnothing \subseteq T$。集合 S 是这个偏序集的最大元素，因为只要 T 是 S 的子集，就有 $T \subseteq S$。■

例 16 在偏序集$(\mathbf{Z}^+, \mid)$中是否存在最大元素和最小元素？

解 1 是最小元素，因为只要 n 是正整数，就有 $1 \mid n$。因为没有被所有正整数整除的整数，所以不存在最大元素。■

有时候可以找到一个元素大于偏序集$(S, \preccurlyeq)$的子集 A 中所有的元素。如果 u 是 S 的元素使得对所有的元素 $a \in A$，有 $a \preccurlyeq u$，那么 u 叫做 A 的一个**上界**。类似地，也可能存在一个元素小于 A 中的所有其他元素。如果 l 是 S 的一个元素使得对所有的元素 $a \in A$ 有 $l \preccurlyeq a$，那么 l 叫做 A 的一个**下界**。

例 17　找出在图 7-20 所示哈斯图的偏序集的子集$\{a, b, c\}$，$\{j, h\}$和$\{a, c, d, f\}$的下界和上界。

解　$\{a, b, c\}$的上界是 e，f，j 和 h，它的唯一的下界是 a。$\{j, h\}$没有上界，它的下界是 a，b，c，d，e 和 f。$\{a, c, d, f\}$的上界是 f，h 和 j，它的下界是 a。■

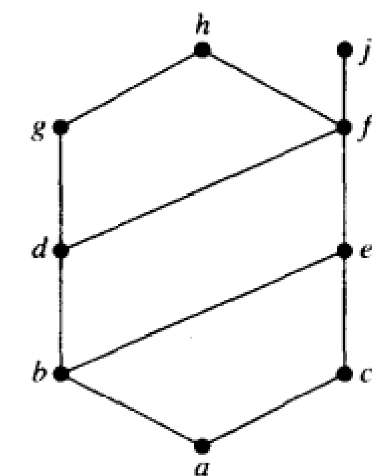


图 7-20　偏序集的哈塞图

元素 x 叫作子集 A 的最小上界，如果 x 是一个上界并且它小于 A 的其他上界。因为如果这样的元素存在，只存在一个，称这个元素为最小上界是有意义的[见节末练习 36(a)]。即如果只要 $a \in A$ 就有 $a \preccurlyeq x$，并且只要 z 是 A 的上界，就有 $x \preccurlyeq z$，x 就是 A 的最小上界。类似地，如果 y 是 A 的下界，并且只要 z 是 A 的下界，就有 $z \preccurlyeq y$，y 就是 A 的最大下界。A 的最大下界如果存在也是唯一的[见节末的练习 36(b)]。一个子集 A 的最大下界和最小上界分别记作 glb(A)和 lub(A)。■

例 18　在图 7-20 所示偏序集中如果$\{b, d, g\}$的最大下界和最小上界存在，求出这个最大下界和最小上界。

解　$\{b, d, g\}$的上界是 g 和 h。因为 $g \prec h$，g 是最小上界。$\{b, d, g\}$的下界是 a 和 b。因为 $a \prec b$，b 是最大下界。■

例 19　在偏序集($\mathbf{Z}^+$，|)中，如果集合{3，9，12}和{1，2，4，5，10}的最大下界和最小上界存在，求出这些最大下界和最小上界。

解　如果 3，9，12 被一个整数整除，那么这个整数就是{3，9，12}的下界。这样的整数只有 1 和 3。因为 1 | 3，3 是{3，9，12}的最大下界。集合{1，2，4，5，10}关系到 | 的下界只有 1，因此 1 是{1，2，4，5，10}的最大下界。

一个整数是{3，9，12}的上界，当且仅当它被 3，9 和 12 整除。具有这种性质的整数就是那些被 3，9 和 12 的最小公倍数 36 整除的整数。因此，36 是{3，9，12}的最小上界。一个正整数是集合{1，2，4，5，10}的上界，当且仅当它被 1，2，4，5 和 10 整除。具有这种性质的整数就是被这些整数的最小公倍数 20 整除的整数。因此，20 是{1，2，4，5，10}的最小上界。■

7.6.5　格

如果一个偏序集的每对元素都有最小上界和最大下界，就称这个偏序集为格。格有许多特殊的性质。此外，格有许多不同的应用，如用在信息流的模型。格在布尔代数中也起到了重要的作用。

例 20　确定图 7-21 的每个哈塞图表示的偏序集是否是格。

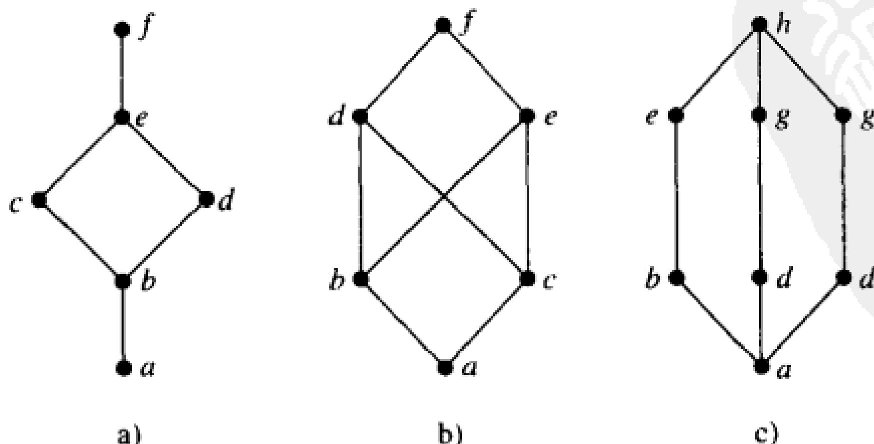


图 7-21　三个偏序集的哈塞图

解 在图7-21a和图7-21c中的哈塞图表示的偏序集是格，因为在每个偏序集中每对元素都有最小上界和最大下界，读者应该能验证这一点。另一方面，图7-21b所示的哈斯图的偏序集不是格，因为元素 b 和 c 没有最小上界。为此只要注意到 d，e 和 f 中每一个都是上界，但这3个元素的任何一个关于这个偏序集中的序都不大于其他2个。■

例21 偏序集$(\mathbf{Z}^+, \mid)$是格吗？

解 设 a 和 b 是两个正整数。这两个整数的最小上界和最大下界分别是它们的最小公倍数和最大公约数，读者应能验证这一点。因此这个偏序集是格。■

例22 确定偏序集$(\{1, 2, 3, 4, 5\}, \mid)$和$(\{1, 2, 4, 8, 16\}, \mid)$是否为格。

解 因为2和3在$(\{1, 2, 3, 4, 5\}, \mid)$中没有上界，它们当然没有最小上界。因此第一个偏序集不是格。

第二个偏序集的每两个元素都有最小上界和最大下界。在这个偏序集中两个元素的最小上界是它们中间较大的元素，而两个元素的最大下界是它们中间较小的元素。读者应能验证这一点。因此第二个偏序集是格。■

例23 确定$(P(S), \subseteq)$是否是格，其中 S 是集合。

解 设 A 和 B 是 S 的两个子集。A 和 B 的最小上界和最大下界分别是 $A \cup B$ 和 $A \cap B$，读者可以证明这一点。因此$(P(S), \subseteq)$是格。■

例24 信息流的格模型 在许多设置中，从一个人或计算机程序到另一个人或计算机程序的信息流要受到限制，这可以通过安全权限来实现。我们可以使用格的模型来表示不同的信息流策略。例如，一个通用的信息流策略是用于政府或军事系统中的*多级安全策略*。为每组信息分配一个安全级别，并且每个安全级别用一个对(A, C)表示，其中 A 是权限级别，C 是种类。然后允许人和计算机程序从一个被特别限制的安全类的集合中访问信息。

在美国政府中，使用的典型的权限级别是不保密(0)、秘密(1)、机密(2)和绝密(3)。在安全级别中使用的种类是一个集合的子集，这个集合含有与一个特定行业领域相关的所有的分部，每个分部表示一个指定的对象域。例如，如果分部的集合是{间谍，鼹鼠，双重间谍}，那么存在8个不同的种类，分部集合有8个子集，对于每个子集有一类，例如{间谍，鼹鼠}。

我们可以对安全类排序，规定$(A_1, C_1) \preccurlyeq (A_2, C_2)$当且仅当 $A_1 \leqslant A_2$ 和 $C_1 \subseteq C_2$。信息允许从安全类(A_1, C_1)流向安全类(A_2, C_2)当且仅当$(A_1, C_1) \preccurlyeq (A_2, C_2)$。例如，信息允许从安全类(机密，{间谍，鼹鼠})流向安全类(绝密，{间谍，鼹鼠，双重间谍})，反之，信息不允许从安全类(绝密，{间谍，鼹鼠})流向安全类(机密，{间谍，鼹鼠，双重间谍})或(绝密，{间谍})。

我们留给读者(见节末练习42)证明所有安全类的集合与在这个例子中所定义的序构成一个格。■

7.6.6 拓扑排序

假设一个项目由20个任务构成。某些任务只能在其他任务结束之后完成。如何找到关于这些任务的顺序？为了对这个问题构造模型，我们建立一个任务集合上的偏序，使得 $a \prec b$，当且仅当 a 和 b 是任务且直到 a 结束后 b 才能开始。为安排好这个项目，需要得出与这个偏序相容的所有20个任务的顺序。我们将说明怎样做到这一点。

我们从定义开始。如果只要 aRb 就有 $a \preccurlyeq b$，则称一个全序 $\preccurlyeq$ 与偏序 R 是*相容的*。从一个偏序构造一个相容的全序叫作*拓扑排序*。我们需要使用引理1。

引理1 每个有穷非空偏序集$(S, \preccurlyeq)$都有极小元素。

证 选择 S 的一个元素 a_0。如果 a_0 不是极小元素，那么存在元素 a_1，满足 $a_1 \prec a_0$。如果 a_1

不是极小元素，那么存在元素 a_2，满足 $a_2 \prec a_1$。继续这一过程，使得如果 a_n 不是极小元素，那么存在元素 a_{n+1} 满足 $a_{n+1} \prec a_n$。因为在这个偏序集只有有穷个元素，这个过程一定结束并且具有极小元素 a_n。 □

我们将要描述的拓扑排序算法对任何有穷非空偏序集都有效。为在偏序集$(A, \preccurlyeq)$上定义一个全序，首先选择一个极小元素 a_1；由引理 1 这样的元素存在。接着，正如读者应该验证的，$(A-\{a_1\}, \preccurlyeq)$也是一个偏序集。如果它是非空的，选择这个偏序集的一个极小元素 a_2。然后再取走 a_2，如果还有其他的元素留下来，在 $A-\{a_1, a_2\}$中选择一个极小元素 a_3。继续这个过程，只要还有元素留下来，就在 $A-\{a_1, a_2, \cdots, a_k\}$中选择极小元素 a_{k+1}。

因为 A 是有穷集，这个过程一定终止。最终产生一个元素序列 $a_1, a_2, \cdots, a_n$。所需要的全序定义为

$$a_1 \preccurlyeq a_2 \preccurlyeq \cdots \preccurlyeq a_n$$

这个全序是与初始的偏序相容的。为看出这一点，注意如果在初始的偏序中 $b \prec c$，c 在算法的某个阶段被选择为极小元素，则这时 b 已经被移出，否则 c 就不会是极小元素。关于这个拓扑排序算法的伪码在算法 1 给出。

算法 1 拓扑排序

```
procedure topological sort(S: 有穷偏序集)
k: =1
while S≠∅
begin
    a_k: =S 的极小元素{由引理 1 这样的元素一定存在}
    S: =S-{a_k}
    k: =k+1
end{a_1, a_2, …, a_n 是 S 的相容的偏序}
```

例 25 找出对于偏序集$(\{1, 2, 4, 5, 12, 20\}, \mid)$的一个相容的全序。

解 第一步是选择一个极小元素。这个元素一定是 1，因为它是唯一的极小元素。下一步选择$(\{2, 4, 5, 12, 20\}, \mid)$的一个极小元素。在这个偏序集中有两个极小元素，即 2 和 5。我们选择 5。剩下的元素是$\{2, 4, 12, 20\}$。在这一步，唯一极小元素是 2。下一步选择 4，因为它是$(\{4, 12, 20\}, \mid)$的唯一极小元素。因为 12 和 20 都是$(\{12, 20\}, \mid)$的极小元素，下一步选哪一个都可以。我们选 20，只剩下 12 作为最后的元素。这产生了全序

$$1 \prec 5 \prec 2 \prec 4 \prec 20 \prec 12$$

这个排序算法所使用的步骤在图 7-22 中给出。 ■

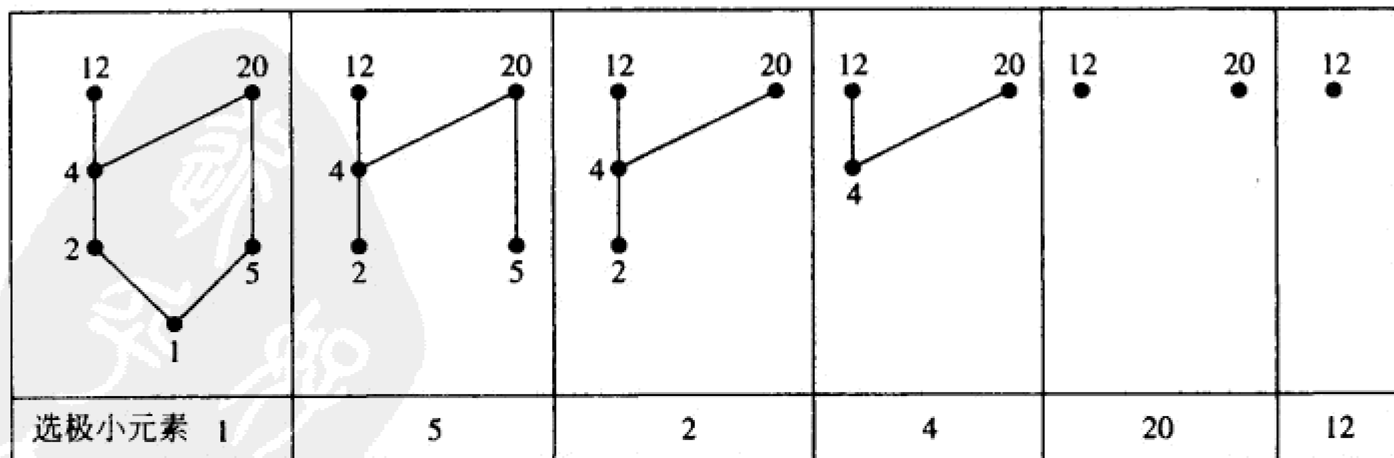


图 7-22 $(\{1, 2, 4, 5, 12, 20\}, \mid)$的拓扑排序

在项目的安排中常会用到拓扑排序。

例 26 一个计算机公司的开发项目需要完成 7 个任务。其中某些任务只能在其他任务结束后才能开始。考虑如下建立任务上的偏序，如果任务 Y 在 X 结束后才能开始，则任务 $X \prec$ 任务 Y。这 7 个任务关于这个偏序的哈塞图在图 7-23 中给出。求一个全序，使得可以按照这个全序执行这些任务以完成这个项目。

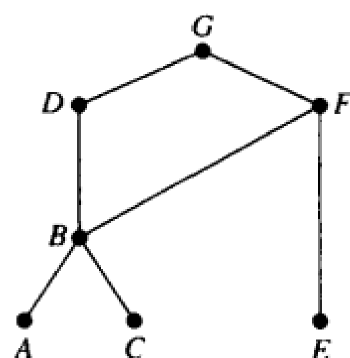


图 7-23 关于 7 个任务的哈塞图

解 可以通过执行一个拓扑排序得到 7 个任务的排序。排序的步骤显示在图 7-24 中。这个排序的结果，$A \prec C \prec B \prec E \prec F \prec D \prec G$，给出了关于任务的一种可能的次序。■

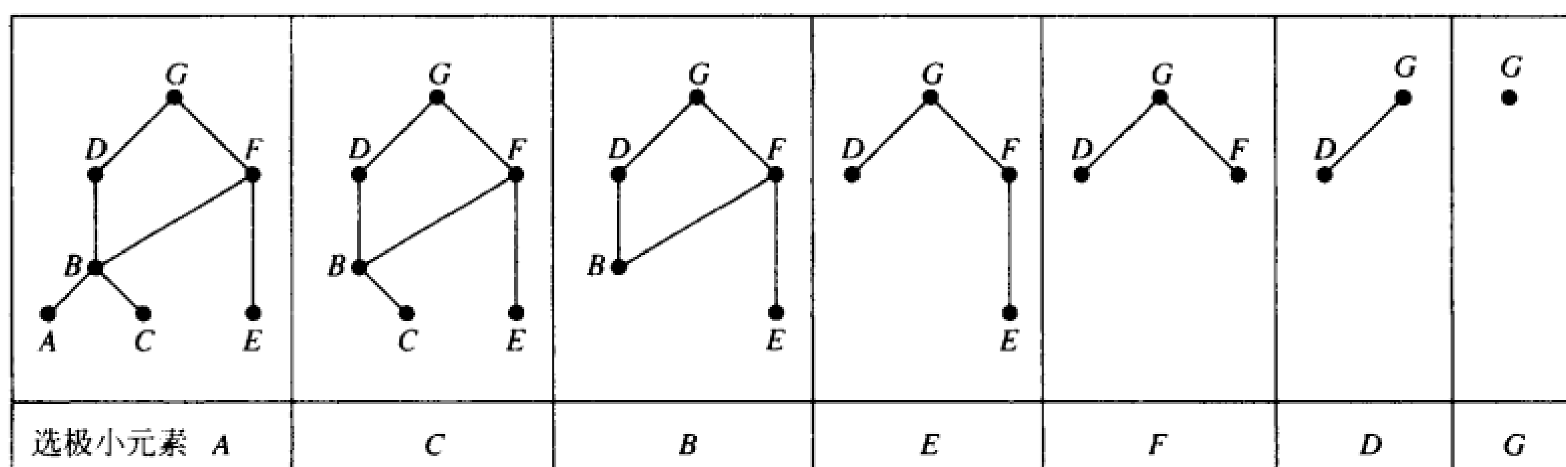


图 7-24 任务的拓扑排序

练习

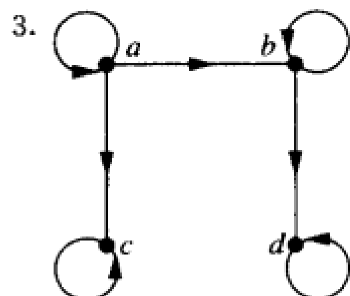
1. 下面哪些集合是偏序集？

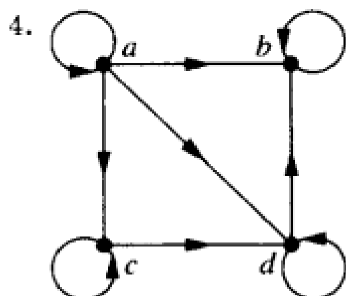
   a)$(\mathbf{Z}, =)$　　b)$(\mathbf{Z}, \neq)$　　c)$(\mathbf{Z}, \geqslant)$　　d)$(\mathbf{Z}, \nmid)$

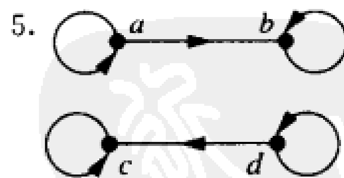
2. 确定由下面的 0－1 矩阵表示的关系是否为偏序。

   a) $\begin{bmatrix}1 & 0 & 1\\ 1 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$　　b) $\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 1 & 0 & 1\end{bmatrix}$　　c) $\begin{bmatrix}1 & 0 & 1 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 1 & 1\\ 1 & 1 & 0 & 1\end{bmatrix}$

在练习 3～5 确定有向图所表示的关系是否为偏序。

3. 

4. 

5. 

6. 设(S, R)是偏序集。证明(S, R^{-1})也是偏序集，其中 R^{-1} 是 R 的逆。偏序集(S, R^{-1})叫做(S, R)的对偶。

7. 求下面偏序集的对偶。

   a)$(\{0, 1, 2\}, \leqslant)$　　b)$(\mathbf{Z}, \geqslant)$　　c)$(P(\mathbf{Z}), \supseteq)$　　d)$(\mathbf{Z}^{+}, \mid)$

8. 在偏序集$(\mathbf{Z}^{+}, \mid)$中下面哪对元素是可比的？

   a)5，15　　b)6，9　　c)8，16　　d)7，7

9. 在下面偏序集中找出两个不可比的元素。

a) $(P(\{0, 1, 2\}), \subseteq)$　　b) $(\{1, 2, 4, 6, 8\}, |)$

10. 设 $S=\{1, 2, 3, 4\}$。考虑基于通常"小于"关系的字典顺序。

a) 找出所有在 $S \times S$ 中小于 $(2, 3)$ 的对。　　b) 找出所有在 $S \times S$ 中大于 $(3, 1)$ 的对。

c) 画出偏序集 $(S \times S, \preccurlyeq)$ 的哈塞图。

11. 找出下面的 n 元组的字典顺序。

a) $(1, 1, 2)$，$(1, 2, 1)$　　b) $(0, 1, 2, 3)$，$(0, 1, 3, 2)$

c) $(1, 0, 1, 0, 1)$，$(0, 1, 1, 1, 0)$

12. 找出下面小写英语字母串的字典顺序。

a) *quack*，*quick*，*quicksilver*，*quicksand*，*quacking*

b) *open*，*opener*，*opera*，*operand*，*opened*

c) *zoo*，*zero*，*zoom*，*zoology*，*zoological*

13. 找出二进位串 0，01，11，001，010，011，0001 和 0101 的基于 $0<1$ 的字典顺序。

14. 画出 $\{0, 1, 2, 3, 4, 5\}$ 上的"大于或等于"关系的哈塞图。

15. 画出 $\{0, 2, 5, 10, 11, 15\}$ 上的"小于或等于"关系的哈塞图。

16. 画出下述集合上的整除关系的哈塞图。

a) $\{1, 2, 3, 4, 5, 6\}$　　b) $\{3, 5, 7, 11, 13, 16, 17\}$

c) $\{2, 3, 5, 10, 11, 15, 25\}$　　d) $\{1, 3, 9, 27, 81, 243\}$

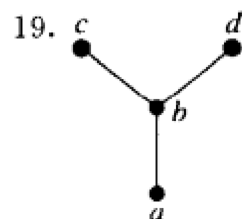
17. 画出下述集合上整除关系的哈塞图。

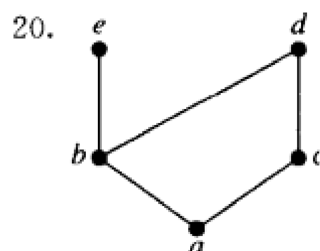
a) $\{1, 2, 3, 4, 5, 6, 7, 8\}$　　b) $\{1, 2, 3, 5, 7, 11, 13\}$

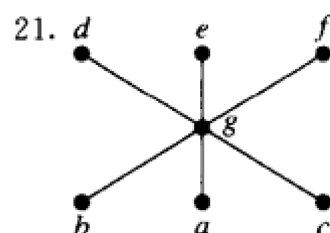
c) $\{1, 2, 3, 6, 12, 24, 36, 48\}$　　d) $\{1, 2, 4, 8, 16, 32, 64\}$

18. 画出在集合 $P(S)$ 上包含关系的哈塞图，其中 $S=\{a, b, c, d\}$。

在练习 19～21 列出所示哈塞图的偏序中的所有有序对。

19. 

20. 

21. 

设 $(S, \preccurlyeq)$ 是偏序集。我们说一个元素 $y \in S$ **覆盖**元素 $x \in S$，当 $x \prec y$ 并且没有元素 $z \in S$ 使得 $x \prec z \prec y$。使得 y 覆盖 x 的对 (x, y) 的集合叫作 $(S, \preccurlyeq)$ 的**覆盖关系**。

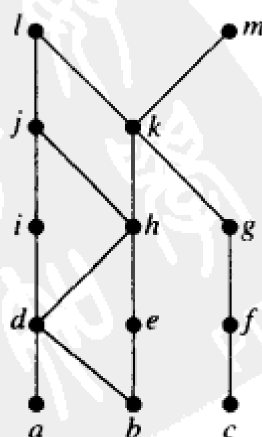
22. $\{1, 2, 3, 4, 6, 12\}$ 上的偏序 $\{(a, b) \mid a$ 整除 $b\}$ 的覆盖关系是什么？

23. S 的幂集上的偏序 $\{(A, B) \mid A \subseteq B\}$ 的覆盖关系是什么？其中 $S=\{a, b, c\}$。

24. 证明有序对 (x, y) 属于有穷偏序集 $(S, \preccurlyeq)$ 的覆盖关系，当且仅当 x 小于 y 并且在这个偏序集的哈塞图中存在一条连接 x 和 y 的边。

25. 证明一个有穷偏序集可以从它的覆盖关系重新构造出来。

26. 对于由右面的哈塞图表示的偏序回答下面的问题。



a) 求极大元素。　　b) 求极小元素。

c) 存在最大元素吗？　　d) 存在最小元素吗？

e) 求 $\{a, b, c\}$ 的所有上界。　　f) 如果存在的话，求 $\{a, b, c\}$ 的最小上界。

g) 求 $\{f, g, h\}$ 的所有下界。　　h) 如果存在的话，求 $\{f, g, h\}$ 的最大下界。

27. 对偏序集 $(\{3, 5, 9, 15, 24, 45\}, |)$，回答下述问题。

a) 求极大元素。　　b) 求极小元素。

c) 存在最大元素吗？　　d) 存在最小元素吗？

e)找出{3，5}的所有上界。　　f)如果存在的话，求{3，5}的最小上界。

g)求{15，45}的所有下界。　　h)如果存在的话，求{15，45}的最大下界。

28. 对偏序集({2，4，6，9，12，18，27，36，48，60，72}，|)，回答下述问题。

a)找出极大元素。　　b)找出极小元素。

c)存在最大元素吗？　　d)存在最小元素吗？

e)找出{2，9}的所有上界。　　f)如果存在，找出{2，9}的最小上界。

g)找出{60，72}的所有下界。　　h)如果存在，找出{60，72}的最大下界。

29. 对偏序集({{1}，{2}，{4}，{1，2}，{1，4}，{2，4}，{3，4}，{1，3，4}，{2，3，4}}，⊆)，回答下述问题。

a)求极大元素。　　b)求极小元素。

c)存在最大元素吗？　　d)存在最小元素吗？

e)求{{2}，{4}}的所有上界。　　f)如果存在的话，求{{2}，{4}}的最小上界。

g)求{{1，3，4}，{2，3，4}}的所有下界。

h)如果存在的话，求{{1，3，4}，{2，3，4}}的最大下界。

30. 给出满足下述条件的偏序集：

a)有一个极小元素但没有极大元素。

b)有一个极大元素但没有极小元素。

c)既没有极大元素也没有极小元素。

31. 证明字典顺序是两个偏序集的笛卡儿积上的偏序。

32. 证明字典顺序是一个偏序集的串的集合上的偏序。

33. 假设$(S, \preccurlyeq_1)$和$(T, \preccurlyeq_2)$是偏序集。证明$(S\times T, \preccurlyeq)$也是偏序集，其中$(s, t)\preccurlyeq(u, v)$，当且仅当$s\preccurlyeq_1 u$且$t\preccurlyeq_2 v$。

34. a)如果偏序集存在最大元素，证明恰好存在一个最大元素。

b)如果偏序集存在最小元素，证明恰好存在一个最小元素。

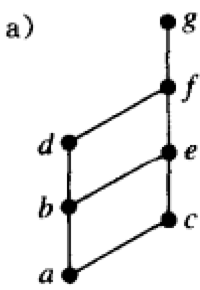
35. a)证明在一个具有最大元素的偏序集中恰好存在一个极大元素。

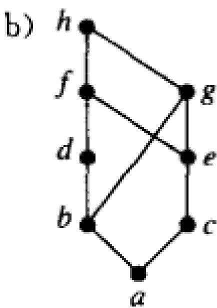
b)证明在一个具有最小元素的偏序集中恰好存在一个极小元素。

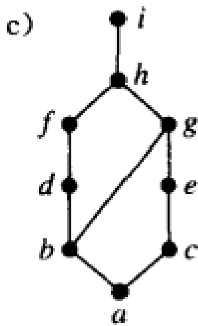
36. a)证明如果偏序集的子集存在最小上界的话，则是唯一的。

b)证明如果偏序集的子集存在最大下界的话，则是唯一的。

37. 确定具有下面哈塞图的偏序集是否为格。

a)


b)


c)


38. 确定下面的偏序集是否为格。

a)({1，3，6，9，12}，|)　　b)({1，5，25，125}，|)

c)$(\mathbf{Z}, \geqslant)$　　d)$(P(S), \supseteq)$，其中$P(S)$是集合S的幂集

39. 证明一个格的每个有限非空子集有最小上界和最大下界。

40. 证明如果偏序集(S, R)是格，那么对偶偏序集(S, R^{-1})也是格。

41. 在一个公司里用信息流的格模型控制敏感信息，这些信息具有由有序对(A, C)表示的安全类别。这里A是权限级别，这种权限级别可以是非私有的(0)、私有的(1)、受限制的(2)或注册的(3)。种类C是所有项目集合{猎豹，黑斑羚，美洲狮}的子集(在公司里常常使用动物的名字作为项目的代码名字)。

a)信息是否允许从(私有的，{猎豹，美洲狮})流向(受限制的，{美洲狮})？

b)信息是否允许从(受限制的，{猎豹})流向(注册的，{猎豹，黑斑羚})？

c)允许信息从(私有的，{猎豹，美洲狮})流向哪个类？

d)允许信息从哪个类流向安全类(受限制的，{黑斑羚，美洲狮})？

42. 证明安全类别(A, C)的集合 S 是一个格，其中 A 是表示权限级别的正整数且 C 是种类的有穷集的子集，具有$(A_1, C_1)\preccurlyeq(A_2, C_2)$，当且仅当 $A_1\leqslant A_2$ 且 $C_1\subseteq C_2$。[提示：首先证明$(S, \preccurlyeq)$是一个偏序集，然后证明(A_1, C_1)和(A_2, C_2)的最小上界和最大下界分别是$(\max(A_1, A_2), C_1\cup C_2)$和$(\min(A_1, A_2), C_1\cap C_2)$。]

*43. 证明一个集合上的所有划分的集合与关系$\preccurlyeq$构成一个格，其中如果划分 P_1 是划分 P_2 的加细，则 $P_1\preccurlyeq P_2$(见 7.5 节的练习 37 前面的说明)。

44. 证明每个全序集都是一个格。

45. 证明每个有限格都有一个最小元素和一个最大元素。

46. 给出一个无限格的例子使得

a)既没有最小元素也没有最大元素。　　b)有一个最小元素但没有最大元素。

c)有一个最大元素但没有最小元素。　　d)有一个最小元素也有一个最大元素。

47. 验证$(\mathbf{Z}^+\times\mathbf{Z}^+, \preccurlyeq)$是一个良序集，其中$\preccurlyeq$是如例 7 所声明的字典顺序。

48. 确定下述每个偏序集是否为良序的。

a)$(S, \leqslant)$，其中 $S=\{10, 11, 12, \cdots\}$

b)$(\mathbf{Q}\cap[0, 1], \leqslant)$(0 和 1 之间的有理数集合，包含 0 和 1)

c)$(S, \leqslant)$，S 是分母不超过 3 的正有理数集合

d)$(\mathbf{Z}^-, \geqslant)$，其中 $\mathbf{Z}^-$ 是负整数的集合

如果偏序集中不存在元素的无限递减序列，即元素 $x_1, x_2, \cdots, x_n\cdots$ 使得 $\cdots\prec x_n\prec\cdots\prec x_2\prec x_1$，则这个偏序集是*良基的*。设偏序集$(R, \preccurlyeq)$，如果对所有的 $x\in S$ 和 $y\in S(x\prec y)$都存在元素 $z\in R$ 使得 $x\prec z\prec y$，则称这个偏序集是*稠密的*。

49. 证明偏序集$(\mathbf{Z}, \preccurlyeq)$(其中 $x\prec y$ 当且仅当 $|x|<|y|$)是良基的，但不是全序集。

50. 证明至少有两个相关元素的稠密的偏序集不是良基的。

51. 证明有理数和通常的“小于或等于”关系构成的偏序集$(\mathbf{Q}, \leqslant)$是稠密偏序集。

*52. 证明小写英文字母的串的集合关于字典顺序既不是良基的，也不是稠密的。

53. 证明偏序集是良序的，当且仅当它是全序的并且是良基的。

54. 证明一个有穷非空偏序集有一个极大元素。

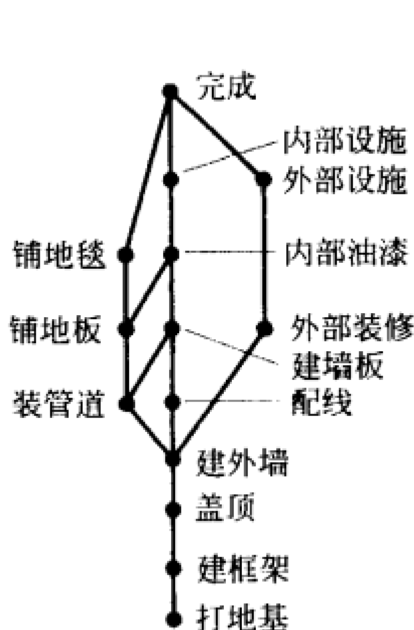
55. 求一个全序，使得它与练习 26 哈塞图所表示的偏序集相容。

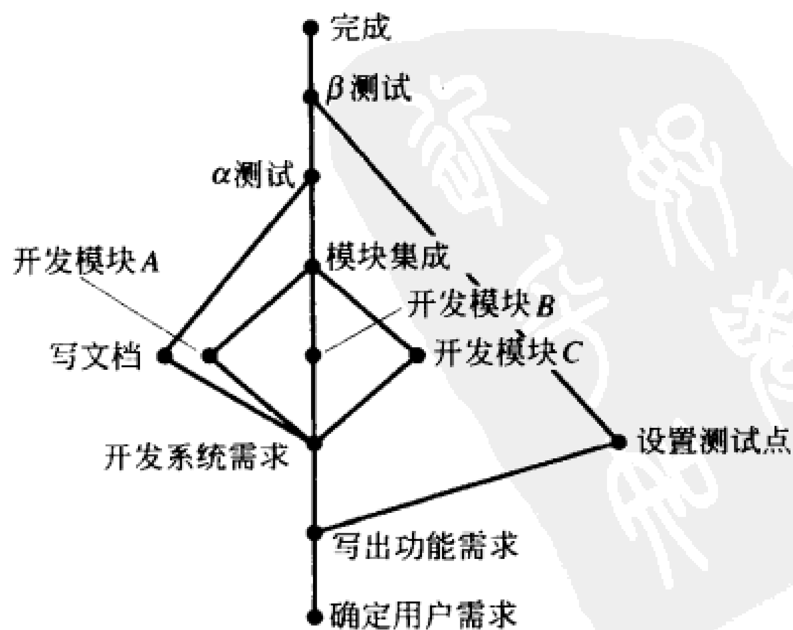
56. 求一个与集合{1，2，3，6，8，12，24，36}上的整除关系相容的全序。

57. 求一个与例 26 构造的序不同的完成开发项目中任务的序。

58. 如果表示建筑一座房子所需任务的哈塞图如下图所示，通过指定这些任务的顺序来安排它们。

59. 如果关于一个软件项目的任务的哈塞图如下所示，对这个软件项目的任务进行排序。





关键术语和结果

术语

从 A 到 B 的二元关系：$A\times B$ 的子集。

A 的关系：从 A 到自身的二元关系(即 $A\times A$ 的子集)。

$S\circ R$：R 与 S 的合成。

R^{-1}：R 的逆关系。

R^n：R 的 n 次幂。

自反的：A 上的一个关系是自反的，如果对所有的 $a\in A$ 有 $(a, a)\in R$。

对称的：A 上的一个关系是对称的，如果只要 $(a, b)\in R$ 就有 $(b, a)\in R$。

反对称的：A 上的关系是反对称的，如果只要 $(a, b)\in R$ 和 $(b, a)\in R$ 就有 $a=b$。

传递的：A 上的关系 R 是传递的，如果从 $(a, b)\in R$ 和 $(b, c)\in R$ 推出 $(a, c)\in R$。

$A_1, A_2, \cdots, A_n$ 的 n 元关系：$A_1\times A_2\times\cdots\times A_n$ 的子集。

关系数据模型：一个使用 n 元关系表示数据库的模型。

主关键字：n 元关系的一个域，使得一个 n 元组被它在这个域的值唯一确定。

复合关键字：一个 n 元关系域的笛卡儿积，使得一个 n 元组被它在这些域的值唯一确定。

选择算子：一个函数，它在 n 元关系中选出满足特定条件的 n 元组。

投影：一个函数，它从一个 n 元关系通过删除域产生一个阶较小的关系。

连接：一个函数，它把具有某些相同的域的 n 元关系组合起来。

有向图：称为顶点的元素以及这些元素的有序对(也叫做边)的集合。

环：形如 (a, a) 的边。

关系 R 关于性质 **P** 的闭包：包含 R 的关系 S(如果存在的话)具有性质 **P**，并且被任何包含 R 且具有性质 **P** 的关系所包含。

有向图中的路径：边的序列 (a, x_1)，(x_1, x_2)，$\cdots$，(x_{n-2}, x_{n-1})，(x_{n-1}, b) 使得序列中每条边的终点是后面一条边的起点。

有向图的回路(或圈)：以同一顶点作为起点和终点的路径。

R^*(连通性关系)：由有序对 (a, b) 构成的关系，条件是存在一条从 a 到 b 的路径。

等价关系：自反的、对称的和传递的关系。

等价：如果 R 是等价关系，若 aRb，那么 a 等价于 b。

$[a]_R$(a 关于 R 的等价类)：A 中所有等价于 a 的元素的集合。

$[a]_m$(模 m 的同余类)：与 a 模 m 同余的整数的集合。

集合 S 的划分：一族两两不相交的非空子集，且这些子集的并就是 S。

偏序：自反的、反对称的和传递的关系。

偏序集 (S, R)：集合 S 与这个集合上的偏序 R。

可比的：偏序集 $(A, \preccurlyeq)$ 的元素 a 和 b 是可比的如果 $a\preccurlyeq b$ 或 $b\preccurlyeq a$。

不可比的：一个偏序集的元素不是可比的。

全序(或线序)：一个偏序，并且它的每对元素都是可比的。

全序(或线序)集：具有一个全序(或线序)的偏序集。

良序集：偏序 $(S, \preccurlyeq)$，其中 $\preccurlyeq$ 是全序，S 的每个非空子集至少有一个元素。

字典顺序：笛卡儿积或串的一个偏序。(参见7.6.2节)

哈塞图：偏序集的图表示，其中所有的环和由传递性导出的边不出现，并且顶点的位置指示了边的方向。

极大元素：偏序集的一个元素，它不小于这个偏序集的任何其他元素。

极小元素：偏序集的一个元素，它不大于这个偏序集的任何其他元素。

最大元素：偏序集的一个元素，它大于或等于这个集合的所有其他元素。

最小元素：偏序集的一个元素，它小于或等于这个集合的所有其他元素。

集合的上界：偏序集的一个元素，它大于这个集合的所有其他元素。

集合的下界：偏序集的一个元素，它小于这个集合的所有其他元素。

集合的最小上界：集合的一个上界，它小于所有其他的上界。

集合的最大下界：集合的一个下界，它大于所有其他的下界。

格：一个偏序集，其中每对元素都有一个最大下界和一个最小上界。

与一个偏序相容的全序：包含了给定偏序的一个全序。

拓扑排序：用给定偏序构造一个相容的全序。

结果

- 集合 A 的关系 R 的自反闭包等于 $R\cup\Delta$，其中 $\Delta=\{(a,a)\mid a\in A\}$。
- 集合 A 的关系 R 的对称闭包等于 $R\cup R^{-1}$，其中 $R^{-1}=\{(b,a)\mid (a,b)\in R\}$。
- 一个关系的传递闭包等于从这个关系构成的连通性关系。
- 为了求一个关系的传递闭包的沃舍尔算法(见 7.4.5 节)。
- 设 R 是等价关系，那么下面三条语句是等价的：(1)aRb；(2)$[a]_R\cap[b]_R\neq\varnothing$；(3)$[a]_R=[b]_R$。
- 集合 A 的等价关系的等价类构成 A 的划分。相反，从一个划分可以构造一个等价关系使得等价类就是划分中的子集。
- 良序归纳原理。
- 拓扑排序算法(见 7.6.5 节)。

复习题

1. a)什么是集合上的关系？　　b)一个 n 元素集合上有多少个关系？
2. a)什么是自反关系？　　b)什么是对称关系？
   c)什么是反对称关系？　　d)什么是传递关系？
3. 给出一个集合{1，2，3，4}上的关系的例子，使得它是
   a)自反的、对称的但不是传递的。　　b)不是自反的，是对称的和传递的。
   c)自反的、反对称的，但不是传递的。　　d)自反的、对称的和传递的。
   e)自反的、反对称的和传递的。
4. a)在一个 n 元素集合上有多少个自反的关系？
   b)在一个 n 元素集合上有多少个对称的关系？
   c)在一个 n 元素集合上有多少个反对称的关系？
5. a)解释在一个大学里怎样用一个 n 元关系表示有关学生的信息。
   b)怎样用一个包含学生姓名、地址、电话号码、专业和平均学分绩点的 5 元关系构造包含学生姓名、专业和平均学分绩点的 3 元关系。
   c)怎样将包含学生姓名、地址、电话号码和专业的 4 元关系和包含学生姓名、学号、专业和学分数的 4 元关系组合成一个单一的 n 元关系？
6. a)解释怎样使用一个 0－1 矩阵表示有穷集上的关系。

b)解释怎样使用表示关系的0－1矩阵来确定这个关系是否为自反的、对称的和反对称的。

7. a)解释怎样使用一个有向图来表示有穷集上的关系。

   b)解释怎样使用表示关系的有向图来确定这个关系是否为自反的、对称的和反对称的。

8. a)定义一个关系的自反闭包和对称闭包。

   b)可以怎样构造一个关系的自反闭包？

   c)可以怎样构造一个关系的对称闭包？

   d)求集合{1，2，3，4}上的关系{(1，2)，(2，3)，(2，4)，(3，1)}的自反闭包和对称闭包。

9. a)定义一个关系的传递闭包。

   b)一个关系的传递闭包能够通过包含所有的对(a, c)得到吗？其中(a, c)满足(a, b)和(b, c)属于R。

   c)描述求关系的传递闭包的两个算法。

   d)求关系{(1，1)，(1，3)，(2，1)，(2，3)，(2，4)，(3，2)，(3，4)，(4，1)}的传递闭包。

10. a)定义等价关系。

    b)集合$\{a, b, c, d\}$上的哪些关系是包含了(a, b)和(b, d)的等价关系？

11. a)证明模m同余关系是等价关系，其中m是正整数。

    b)证明关系$\{(a, b) \mid a \equiv \pm b (\mathrm{mod} 7)\}$是整数集上的等价关系。

12. a)什么是一个等价关系的等价类？

    b)什么是模5同余关系的等价类？

    c)什么是问题11(b)中等价关系的等价类？

13. 解释在集合的等价关系与集合划分之间的联系。

14. a)定义偏序。 b)证明正整数集合上的整除关系是偏序。

15. 解释怎样用集合A_1和A_2上的偏序定义在集合$A_1 \times A_2$上的偏序。

16. a)解释怎样构造有穷集上的偏序的哈塞图。

    b)画出集合{2，3，5，9，12，15，18}上的整除关系的哈塞图。

17. a)定义一个偏序集的极大元素和最大元素。

    b)给出一个有3个极大元素的偏序集的例子。

    c)给出一个有1个最大元素的偏序集的例子。

18. a)定义格。

    b)给出一个5元偏序集是格的例子和一个5元偏序集不是格的例子。

19. a)证明一个格的每个有穷子集有一个最大下界和一个最小上界。

    b)证明每个具有有限元素的格有一个最小元素和一个最大元素。

20. a)定义一个良序集。

    b)描述一个从偏序集产生良序集的算法。

    c)如果仅当某个或某些其他任务完成以后，一个任务才可以开始，解释怎样用(b)中的算法排序这个项目中的任务。

补充练习

1. 设S是所有英语字母的串的集合。确定下面的关系是否是自反的、反自反的、对称的、反对称的和传递的。

   a)$R_1 = \{(a, b) \mid a$和b没有公共字母$\}$

   b)$R_2 = \{(a, b) \mid a$和b长度不相等$\}$

   c)$R_3 = \{(a, b) \mid a$比b长$\}$

2. 构造集合$\{a, b, c, d\}$上的关系，使得它是

   a)自反的、对称的，但不是传递的。

b)反自反的、对称的和传递的。

c)反自反的、反对称的，但不是传递的。

d)自反的，既不是对称的也不是反对称的，是传递的。

e)既不是自反的、反自反的、对称的和反对称的，也不是传递的。

3. $\mathbf{Z}\times\mathbf{Z}$ 上的关系 R 定义如下：$(a, b)R(c, d)$，当且仅当 $a+d=b+c$。证明 R 是等价关系。
4. 证明一个反对称关系的子集也是一个反对称关系。
5. 设 R 是 A 上的自反关系，证明 $R\subseteq R^2$。
6. 假设 R_1 和 R_2 是集合 A 上的自反关系，证明 $R_1\oplus R_2$ 是反自反的。
7. 假设 R_1 和 R_2 是集合 A 上的自反关系，$R_1\cap R_2$ 也是自反的吗？$R_1\cup R_2$ 也是自反的吗？
8. 假设 R 是集合 A 上的对称关系，$\bar{R}$ 也是对称的吗？
9. 设 R_1 和 R_2 是集合 A 上的对称关系，$R_1\cap R_2$ 也是对称的吗？$R_1\cup R_2$ 也是对称的吗？
10. 一个关系 R 叫作*循环的*，如果 aRb 和 bRc 推出 cRa。证明 R 是自反的和循环的，当且仅当它是等价关系。
11. 证明一个 n 元关系的主关键字也是这个关系的任何投影的主关键字，其中这个投影包含这个关键字作为它的一个域。
12. 一个 n 元关系的主关键字也就是由取这个关系与第二个关系的连接而得到的较大的关系的主关键字吗？
13. 证明一个关系的对称闭包的自反闭包和它的自反闭包的对称闭包是相同的。
14. 设 R 是所有数学家的集合上的关系，R 包含有序对 (a, b)，当且仅当 a 与 b 合写了一篇论文。

    a)描述关系 R^2。　　b)描述关系 R^*。

    c)如果一个数学家与多产的匈牙利数学家保罗·埃德斯[⊖]合写了一篇论文，那么这个数学家的*埃德斯数*是 1。如果这个数学家没有与埃德斯合写一篇论文，但是与某个与埃德斯合写过论文的人合写了一篇论文，那么这个数学家的埃德斯数是 2，依此下去(埃德斯本人的埃德斯数是 0)。用 R 中路径的概念给出埃德斯数的定义。
15. a)给出一个例子，证明一个关系的对称闭包的传递闭包不一定与这个关系的传递闭包的对称闭包相等。

    b)证明一个关系的对称闭包的传递闭包一定包含这个关系的传递闭包的对称闭包。
16. a)设 S 是一个计算机问题的子程序的集合。定义关系 R，如果在执行中子程序 **P** 调用子程序 **Q**，那么 **P** R **Q**。描述 R 的传递闭包。

⊖ **保罗·埃德斯(Paul Erdös, 1913—1996)**　埃德斯生于匈牙利的布达佩斯，是两个高中数学教师的儿子。他是一个神童：3 岁时就能心算 3 位数的乘法，4 岁时独自发现了负数。由于母亲不想让他外出染上传染病，他主要在家读书。17 岁时埃德斯进入 Eötvös 大学，4 年后获得数学博士毕业。之后，在英国曼彻斯特做了 4 年的博士后。1938 年由于匈牙利的政治形势，特别是对犹太人的形势恶化，他去了美国。他在美国度过了一生的大部分时光。1954 年至 1962 年，由于麦卡锡的偏执政策，他被驱逐出美国。他还在以色列度过了相当长的时间。

埃德斯对组合学和数论做出了许多突出的贡献。他的最了不起的发现之一是关于素数定理的基础证明(这里的含义是指没有使用任何复杂的分析)。这个定理对于不超过一个固定正整数的素数的个数进行了估计。他也为拉姆齐理论的近代发展做出了贡献。

埃德斯周游世界，出席会议，访问大学和研究所，与其他的数学家一起工作。他全身心地投身数学，不间断地寻访各个数学家，并宣称“我的大脑是开放的”。埃德斯著述和合写的论文有 1500 多篇，论文的合作者有 500 多个。因为他没有固定的家，这些论文的备份由 AT&T 实验室著名的离散数学家劳恩·格雷汉姆(Ron Graham)保存，这个数学家与他广泛合作，并在生活上给予他很多关照。

埃德斯设立从 10 美元到 10 000 美元的奖金，用于对他特别感兴趣的问题的求解，奖金的多少依赖于问题的难度。他已支付的奖金差不多有 4000 美元。埃德斯有自己的特殊语言，使用的短语有：“epsilon”(孩子)，“boss”(女人)，“slave”(男人)，“captured”(结婚)，“liberated”(离婚)，“Supreme Fascist”(上帝)，“Sam”(美国)和“Joe”(苏联)。尽管他在许多事上与众不同，但是他几乎把所有的精力都投入在数学研究上。他没有嗜好，也没有全日制的工作。他一生未婚，看起来是独身主义者。埃德斯特别慷慨，把得到的许多奖品、奖金和薪金用于学术交流和有价值的事业。他没有很强的物欲，他的旅行十分简易。

b)对于哪些子程序 P，$(\mathbf{P}, \mathbf{P})$属于 R 的传递闭包？

c)描述 R 的传递闭包的自反闭包。

17. 假设 R 和 S 是集合 A 上的关系，$R\subseteq S$，且 R 和 S 的关于性质 $\mathbf{P}$ 的闭包存在。证明 R 关于 $\mathbf{P}$ 的闭包是 S 关于 $\mathbf{P}$ 的闭包的子集。
18. 证明两个关系的并的对称闭包是它们的对称闭包的并。

*19. 设计一个基于内点概念的算法，求有向图中两个顶点之间的最长路径的长度，或确定在这些顶点之间存在任意长的路径。

20. 下面的哪些关系是所有人集合上的等价关系？

a)$\{(x, y) \mid x$ 与 y 有同样的星座$\}$

b)$\{(x, y) \mid x$ 与 y 出生在同一年$\}$

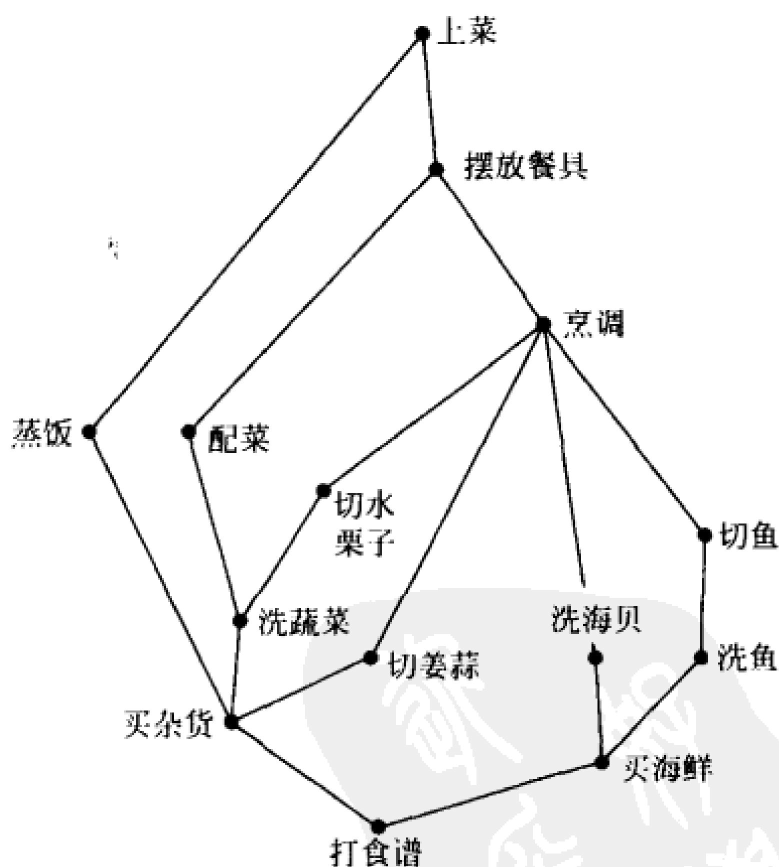
c)$\{(x, y) \mid x$ 与 y 曾待在同一城市$\}$

*21. 在5元素集合上有多少个不同的等价关系恰有3个不同的等价类？

22. 证明$\{(x, y) \mid x-y\in \mathbf{Q}\}$是实数集上的等价关系，其中 $\mathbf{Q}$ 是有理数集合。$[1]$，$[1/2]$和$[\pi]$是什么？
23. 设 $P_1=\{A_1, A_2, \cdots, A_m\}$和 $P_2=\{B_1, B_2, \cdots, B_n\}$都是集合 S 的划分。证明形如 $A_i\cap B_j$ 的非空子集族是 S 的划分，且是 P_1 和 P_2 的加细(见7.5节练习37前面的说明)。

*24. 证明关系 R 的自反闭包的对称闭包的传递闭包是包含 R 的最小的等价关系。

25. 设 $R(S)$是集合 S 上的所有关系的集合。如下定义 $R(S)$上的关系$\preccurlyeq$，如果 $R_1\subseteq R_2$ 则 $R_1\preccurlyeq R_2$，这里的 R_1 和 R_2 是 S 上的关系。证明$(R(S), \preccurlyeq)$是偏序集。
26. 设 $\mathbf{P}(S)$是 S 上的所有划分的集合。定义 $\mathbf{P}(S)$上的关系如下：如果 P_1 是 P_2 的加细，则 $P_1\preccurlyeq P_2$(见7.5节的练习37)。证明$(\mathbf{P}(S), \preccurlyeq)$是偏序集。
27. 下图为烹调中餐任务的哈塞图，通过规定那些必须任务的顺序来安排烹调。



如果一个偏序集子集的每一对元素都是可比的，就称这个子集为一条**链**。如果它的每一对元素都是不可比的，就称这个子集为一条**反链**。

28. 找出7.6节练习19～21的哈塞图所表示的偏序集中所有的链。
29. 找出7.6节练习19～21的哈塞图所表示的偏序集中所有的反链。
30. 找出7.6节练习26的哈塞图所示偏序集中具有最多元素的一条反链。
31. 证明在一个有穷偏序集$(S, \preccurlyeq)$中每个极大链包含 S 的一个极小元素。(一条极大链是一条链，但不是一条更大的链的子链。)

**32. 证明一个偏序集可以被分成 k 条链，其中 k 是在这个偏序集中一条反链的最多元素数。

*33. 证明在任意一组 $mn+1$ 个人中要么存在 $m+1$ 个人的一个列表，其中每个人(除了表的第一个人以外)都是表中前一个人的后代；要么存在 $n+1$ 个人，其中没有一个人是其他 n 个人中任何一个人的后代。[提示：用练习32。]

假设$(S, \preccurlyeq)$是良基的偏序集。**良基归纳原理**说，如果$\forall x\forall y(y\prec x\rightarrow P(y))\rightarrow P(x))$，那么对所有的 $x\in S$，有 $P(x)$为真。

34. 证明对于良基归纳原理不需要另外的基础情况，就是说，如果$\forall x(\forall y(y\prec x\rightarrow P(y))\rightarrow P(x))$、$P(u)$对于 S 中所有的极小元素 u 为真。

*35. 证明良基归纳原理是有效的。

集合 A 上的关系 R 叫做 A 上的近似序，如果 R 是自反的和传递的。

36. 设 R 是从 $\mathbf{Z}^+$ 到 $\mathbf{Z}^+$ 的所有函数的集合上的关系，使得 (f, g) 属于 R，当且仅当 f 是 $O(g)$，证明 R 是近似序。

37. 设 R 是 A 上的近似序，证明 $R\cap R^{-1}$ 是等价关系。

*38. 设 R 是近似序，设 S 是 $R\cap R^{-1}$ 的等价类的集合上的关系，C 和 D 是 R 的等价类，(C, D) 属于 S，当且仅当存在 C 的元素 c，D 的元素 d，使得 (c, d) 属于 R。证明 S 是偏序。

设 L 是一个格。由 $x\wedge y=\text{glb}(x, y)$ 和 $x\vee y=\text{lub}(x, y)$ 定义交($\wedge$)和并($\vee$)运算。

39. 证明下面的性质对格 L 的一切元素 x，y，z 成立

a) $x\wedge y=y\wedge x$，$x\vee y=y\vee x$　（交换律）

b) $(x\wedge y)\wedge z=x\wedge(y\wedge z)$，$(x\vee y)\vee z=x\vee(y\vee z)$　（结合律）

c) $x\wedge(x\vee y)=x$，$x\vee(x\wedge y)=x$　（吸收律）

d) $x\wedge x=x$，$x\vee x=x$　（幂等律）

40. 证明如果 x 和 y 是格的元素，那么 $x\vee y=y$ 当且仅当 $x\wedge y=x$。

一个格 L 是有界的，如果它有一个上界，记作 1，即对所有的 $x\in L$ 有 $x\leqslant 1$；并且有一个下界，记作 0，即对所有的 $x\in L$ 有 $0\leqslant x$。

41. 证明如果 L 是具有上界 1 和下界 0 的有界格，那么对所有的元素 $x\in L$，具有下面的性质：

a) $x\vee 1=1$　　b) $x\wedge 1=x$　　c) $x\vee 0=x$　　d) $x\wedge 0=0$

42. 证明每个有限格是有界的。

一个格叫作分配的，如果对 L 中所有的 x，y，z，有 $x\vee(y\wedge z)=(x\vee y)\wedge(x\vee z)$ 和 $x\wedge(y\vee z)=(x\wedge y)\vee(x\wedge z)$。

*43. 给出一个不是分配格的例子。

44. 证明格 $(P(S), \subseteq)$ 是分配格，其中 $P(S)$ 是有穷集 S 的幂集。

45. 格 $(\mathbf{Z}^+, |)$ 是分配格吗？

有界格 L 的元素 a 关于上界 1 和下界 0 的补元是元素 b，使得 $a\vee b=1$ 和 $a\wedge b=0$。如果一个格的每个元素都有补元，那么这个格叫做有补格。

46. 给出一个有限格的例子，其中至少 1 个元素有多于 1 个的补元且至少 1 个元素没有补元。

47. 证明格 $(P(S), \subseteq)$ 是有补格，其中 $P(S)$ 是有穷集 S 的幂集。

*48. 证明如果 L 是有限分配格，那么 L 的一个元素至多有 1 个补元。

计算机题目

按下述给定的输入和输出写出程序。

1. 给定表示有穷集上关系的矩阵，确定这个关系是否是自反或反自反的。
2. 给定表示有穷集上关系的矩阵，确定这个关系是否是对称或反对称的。
3. 给定表示有穷集上关系的矩阵，确定这个关系是否是传递的。
4. 给定正整数 n，显示一个 n 元素集合上所有的关系。

*5. 给定一个正整数 n，确定 n 元素集合上的传递关系的个数。

*6. 给定一个正整数 n，确定 n 元素集合上的等价关系的个数。

*7. 给定一个正整数 n，显示 n 个最小的正整数集合上的所有等价关系。

8. 给定一个 n 元关系，当某些特定的字段被删除以后求这个关系的投影。

9. 给定一个 m 元关系、一个 n 元关系和一个公共字段的集合，找出这些关系关于这些公共字段的连接。

10. 给定表示一个有穷集上关系的矩阵，求表示这个关系自反闭包的矩阵。

11. 给定表示一个有穷集上关系的矩阵，求表示这个关系对称闭包的矩阵。
12. 给定表示一个有穷集上关系的矩阵，通过计算表示这个关系的矩阵的幂的联合求表示这个关系传递闭包的矩阵。
13. 给定表示一个有穷集上关系的矩阵，使用沃舍尔算法求表示这个关系的传递闭包的矩阵。
14. 给定表示一个有穷集上关系的矩阵，求表示包含这个关系的最小的等价关系的矩阵。
15. 给定一个有穷集上的偏序，使用拓扑排序找出一个与它相容的全序。

计算和研究

使用一个计算程序或你已完成的程序做下面的练习。

1. 显示一个4元素集合上的所有不同的关系。
2. 显示一个6元素集合上的所有不同的自反和对称的关系。
3. 显示一个5元素集合上的所有不同的自反和传递的关系。

*4. 对所有的正整数 n，$n\leqslant 7$，确定在 n 元素集合上有多少个传递关系。

5. 使用对应于某个特定的运输或通信网络的有向链路的关系，或者使用一个随机生成的关系，在至少20个元素的集合上求所用关系的传递闭包。
6. 对于所有不超过20的正整数 n，计算 n 元素集合上的不同的等价关系个数。
7. 显示7元素集合上的所有的等价关系。

*8. 显示5元素集合上的所有的偏序。

*9. 显示5元素集合上的所有的格。

写作题目

用本书以外的资料写成短文回答下列问题。

1. 讨论模糊关系的概念。怎样使用模糊关系？
2. 不限于7.2节所介绍的内容，描述关系数据库的基本原理。关系数据库与其他类型的数据库相比，使用面有多广？
3. 查找沃舍尔和罗伊(*Roy*)的原始论文(法文)，在那篇论文中他们提出了求传递闭包的算法。讨论他们的方法。为什么可以认为沃舍尔算法是被多人独立发现的？
4. 描述怎样用等价类把有理数定义为整数对的类，以及遵照这种方法怎样定义有理数的基本算术运算(见7.5节练习10)。
5. 解释赫尔姆·哈塞怎样使用我们现在称为哈塞图的图示。
6. 描述在计算机操作系统中用来执行信息流策略的某些机制。
7. 讨论计划评审技术(Program Evaluation and Review Technique，PERT)在安排一个大的复杂项目的任务中的应用。PERT的使用面有多广？
8. 讨论关键路径方法(Critical Path Method，CPM)对找出完成项目的最短时间的应用。CPM的使用面有多广？
9. 讨论格中的对偶性的概念。解释怎样用对偶性建立新的结果？
10. 解释模格的意义。描述模格的某些性质，描述模格是怎样在投影几何的研究中产生的。



第8章　图

图论是有着许多现代应用的古老课题。伟大的瑞士数学家列昂哈德·欧拉在18世纪就建立了图论的基本思想。他利用图解决了有名的哥尼斯堡七桥问题，在本章里将要讨论这个问题。

图可用来解决许多领域的问题。例如，可用图来确定平面电路板上的电路能否实现；用图来区别分子式相同但结构不同的两种化合物；用图来研究万维网的结构；用计算机网络的图模型来确定两台计算机是否由通信链路所连接；用边上带权值的图来解决诸如寻找城市交通网中两城之间最短通路的问题；还可用图来安排考试和分配电视台的频道。

8.1　概述

图是由一些顶点和连接这些顶点的一些边所组成的离散结构。根据连接顶点对的边的种类和数目的不同，图有多种不同类型。几乎每一门可以想到的学科，都有用图模型来解决的问题。我们将要给出例子来说明图模型在不同领域的应用。例如，将要说明如何用图来表示生态环境里不同物种的竞争，如何用图来表示在组织里谁影响谁，以及如何用图来表示比赛结果。稍后将要说明如何用图解决许多问题，比如计算航线网里两城之间航班的不同组合的数目，确定是否可能走遍城市里所有街道而不重复，以及求出地图区域着色所需要的颜色数。

8.1.1　图的种类

本节将通过说明如何用每一类型的图为计算机网络建模来介绍不同类型的图。假设网络包括计算机和连接计算机的电话线。用点表示每台计算机的位置，用弧表示每条电话线，如图8-1所示。

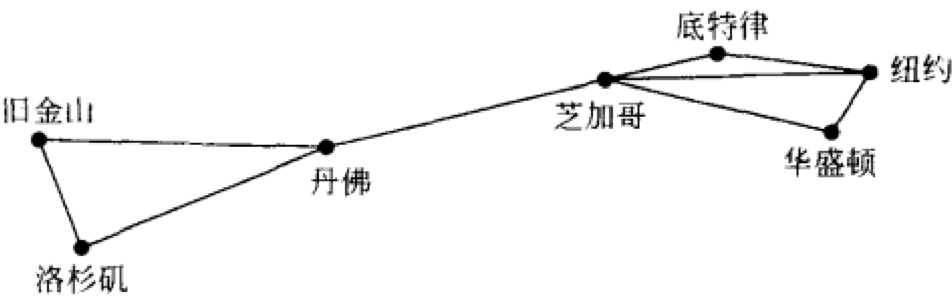


图8-1　计算机网络

在图8-1的网络里，两台计算机之间最多有一条电话线，每条线都是双向传输，没有计算机有到自身的电话线。所以这个网络用简单图建模，由表示计算机的顶点和表示电话线的无向边所组成。每条边都连接两个不同的顶点，而且没有两条边连接同样一对顶点。

定义1　简单图 $G=(V, E)$ 是由非空顶点集 V 和边集 E 所组成的，V 的不同元素的无序对称为边。

当通信量很大时，网络里的计算机之间往往要连接多条电话线，如图8-2所示。简单图已无法为这样的网络建模，代之以多重图，它包括顶点和顶点之间的无向边，允许顶点对之间有多重边。每一个简单图都是多重图。反之，并非所有多重图都是简单图，因为在多重图里，两条或更多条边可以连接同一对顶点。

当出现多重边时，不能用一对顶点来指定图的边。这样使得多重图的定义变得有点复杂。

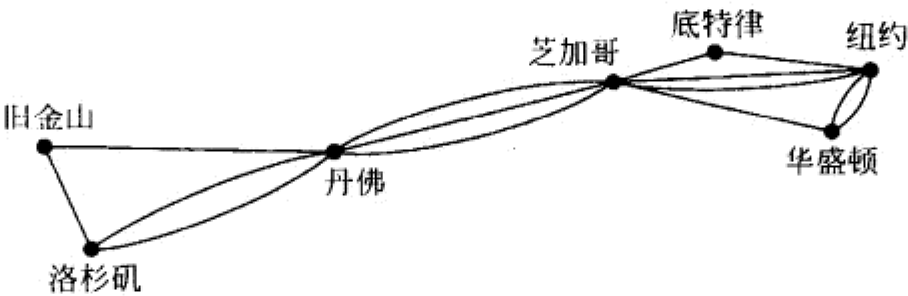


图 8-2 带多重线的计算机网络

定义 2 多重图 $G=(V, E)$是由非空顶点集 V、边集 E 以及从 E 到$\{\{u, v\} \mid u, v\in V, u\neq v\}$的函数 f 所组成的。若 $f(e_1)=f(e_2)$，则边 e_1 和边 e_2 称为多重边或平行边。

读者应当注意在多重图里，多重边关联同一对顶点。不过，若至少有一条边 e 满足 $f(e)=\{u, v\}$，则说$\{u, v\}$是图 $G=(V, E)$的边。不区分边 e 和它所关联的集合$\{u, v\}$，除非识别个别多重边很重要。

计算机网络可以包含计算机连到它自身的电话线(可能是为了诊断的目的)，如图 8-3 所示。不能用多重图为这样的网络建模，因为多重图里不允许有环(即从顶点到它自身的边)，而需要用伪图，伪图比多重图更符合一般情况，因为在伪图里的边可以连接顶点和它自身。

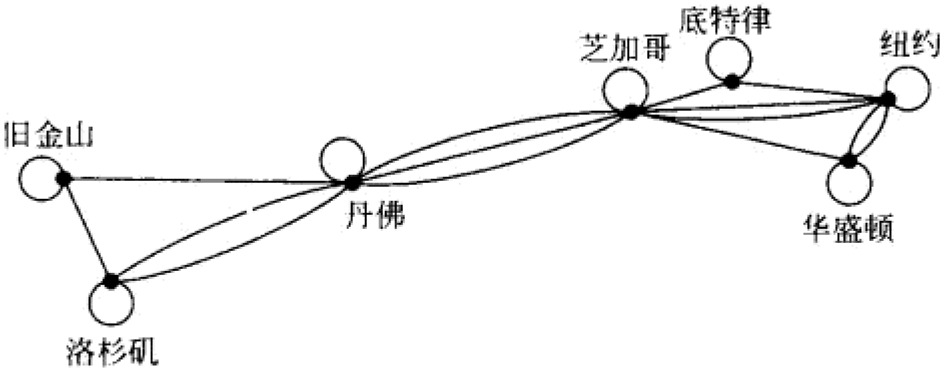


图 8-3 带诊断线的计算机网络

为了形式化地定义伪图，必须设法把边与只包含一个顶点的集合关联起来。

定义 3 伪图 $G=(V, E)$是由非空顶点集 V、边集 E 以及从 E 到$\{\{u, v\} \mid u, v\in V\}$的函数 f 所组成的。若对某个 $u\in V$ 来说有 $f(e)=\{u, u\}=\{u\}$，则边 e 是环。

计算机网络里的电话线可以不是双向传输。例如，图 8-4 里纽约的主机只从其他计算机接收数据而不发出数据。其他电话线用成对的反向边来表示双向传输。

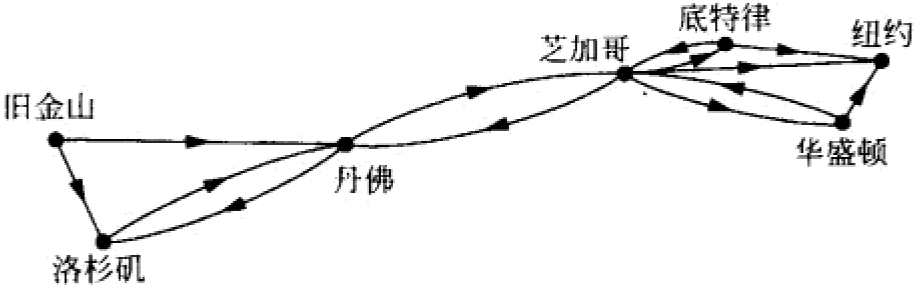


图 8-4 带单向电话线的通信网络

用有向图(在第 7 章里介绍过)为这样的网络建模。有向图的边是有序对。允许有环(即相同元素的有序对)，但不允许在两个顶点之间有同向的多重边。回忆下述的定义。

定义 4 有向图(V, E)是由非空顶点集 V、边集 E 所组成的，边是 V 中元素的有序对。用从 u 指向 v 的箭头表示边(u, v)的方向。

最后，计算机网络里可以出现多重线，所以从每个地点到纽约主机可以有多条单向线，并且从主机回到每个远程计算机有不止一条线，如图 8-5 所示。用有向多重图来为这样的网络建模，它有从一个顶点到第二个(可能相同)顶点的多重有向边。下面给出有向多重图的定义。

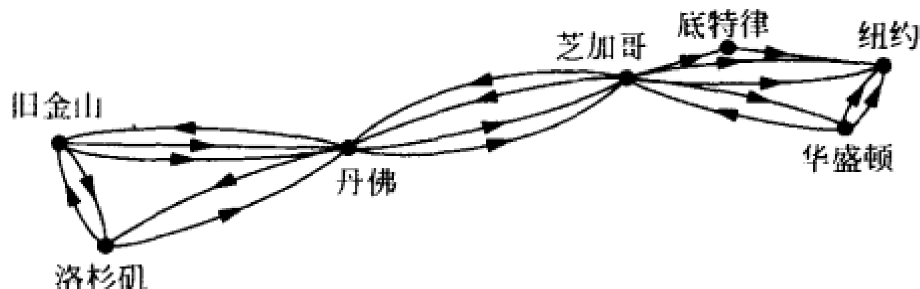


图 8-5 带多重单向线的计算机网络

定义 5 有向多重图 $G=(V, E)$ 是由非空顶点集 V、边集 E 以及从 E 到 $\{(u, v) \mid u, v\in V\}$ 的函数 f 所组成的。若 $f(e_1)=f(e_2)$，则边 e_1 和边 e_2 是多重边。

读者应当注意到，多重有向边所关联的是同一对顶点。不过，只要至少有一条边满足 $f(e)=(u, v)$，就说 (u, v) 是 $G=(V, E)$ 的边。不区分边 e 和它所关联的有序对 (u, v)，除非识别个别多重边很重要。

表 8-1 总结了各种类型的图的术语。将用图来描述带有有向边或无向边、带有或不带有环和多重边的图。名词无向图或伪图表示带有多重边和环的无向图。当图的边包含有序对时，将用有向的来形容。

图论不断发展到现在，应用范围已非常广泛，许多不同的图论术语会经常用到。当遇到这些术语时，读者要理解它们具体的使用意义。

表 8-1 图术语

类　型	边	允许多重边	允许环
简单图	无向	否	否
多重图	无向	是	否
伪图	无向	是	是
有向图	有向	否	是
有向多重图	有向	是	是

8.1.2 图模型

图可用在各种模型里。这里将介绍几种领域的图模型。本章的后续小节和后面几章将要介绍其他模型。

例 1 生态学里栖息地重叠图 图可用在涉及不同种类的动物在一起活动的许多模型里。例如用栖息地重叠图为生态系统里物种之间的竞争建模。用顶点表示每个物种。若两个物种竞争(即它们共享某些食物来源)，则用无向边连接表示它们的顶点。图 8-6 里的图表示森林生态系统。从这个图看出松鼠与浣熊竞争，但是乌鸦不与鼩鼱竞争。■

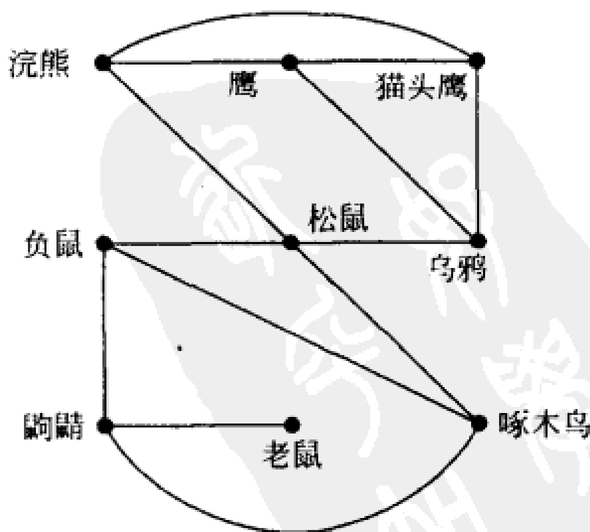


图 8-6 栖息地重叠图

例 2 熟人关系图 可用图模型来表示人与人之间的各种关系。例如可用图来表示两个人是否互相认识，即他们是否是熟人。用顶点表示具体人群里的每个人。当两个人互相认识时，用有向边连接这两个人。图 8-7 显示了一个小型熟人关系图。世界上所有人的熟人关系图有超过 60 亿个顶点和可能超过 1 万亿条边！在 8.4 节里将要进一步讨论这个图。 ■

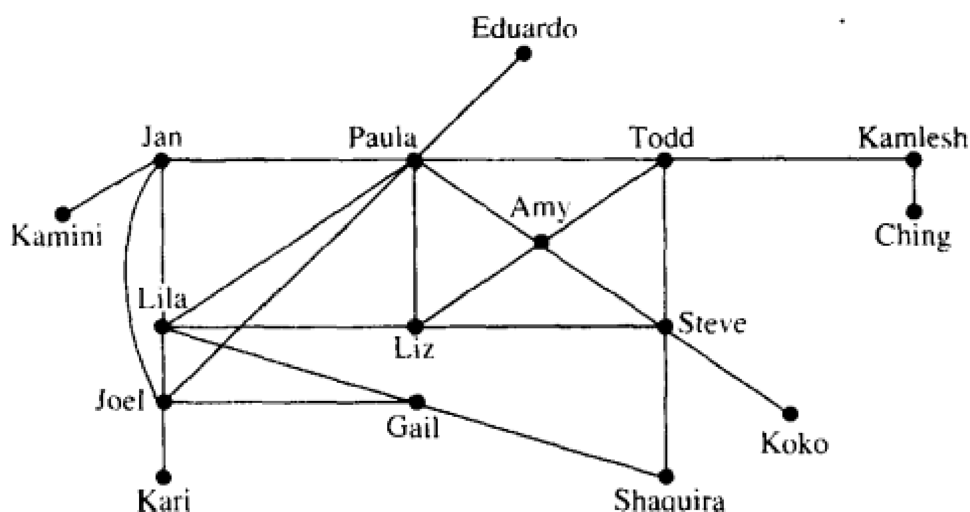


图 8-7 熟人关系图

例 3 影响图 在对群体行为的研究里，可以观察到某些人能够影响其他人的思维。一种称为影响图的有向图可以用来为这样的行为建模。用顶点表示群体的每个人。当顶点 a 所表示的人影响到顶点 b 所表示的人时，就有从顶点 a 到顶点 b 的有向边。图 8-8 表示了群体成员的影响图的例子。在这个用影响图建模的群体里，Deborah 影响 Brain、Fred 以及 Linda，但是没有人影响 Deborah。另外，Yvonne 与 Brain 互相影响。 ■

例 4 好莱坞图 好莱坞图用顶点表示演员，并且当两个顶点所表示的演员共同出演一部电影时，就连接这两个顶点。根据互联网电影数据库，在 2001 年 11 月，好莱坞图有 574 724 个顶点和超过 1600 万条边，这些顶点所表示的演员出现在 292 609 部电影中。稍后在 8.4 节里将要讨论好莱坞图。 ■

例 5 循环赛 每个队都与其他每队恰好比赛一次的联赛称为*循环赛*。用其中顶点表示每个队的有向图来为这样的比赛建模。注意若 a 队击败 b 队，则 (a, b) 是边。图 8-9 表示这样的有向图模型。注意在这个比赛里，队 1 无败绩而队 3 无胜绩。 ■

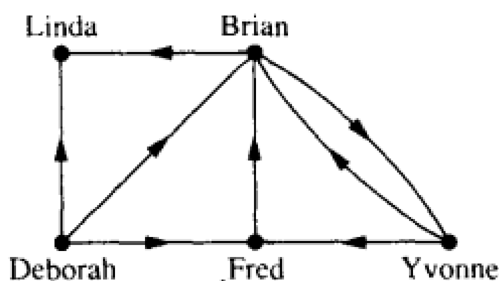


图 8-8 影响图

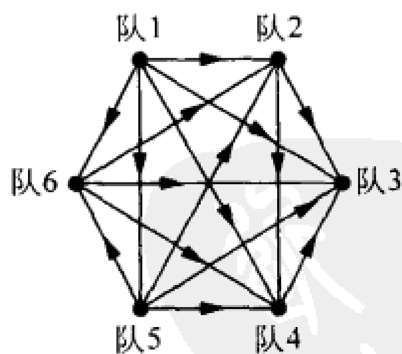


图 8-9 循环赛的图模型

例 6 合作图 合作图可用来为学术论文的合作者关系建模。在合作图中，顶点表示人(可能限制为某个学术圈子的成员)，如果两个人合写论文，则用边连接这两个人。已经发现在数学研究论文上合作者的合作图有超过 337 000 个顶点和 496 000 条边。在 8.4 节里将对这个图补充说明。 ■

例 7 呼叫图 可用来为网络(比如长途电话网)里的电话呼叫建模。具体地说,有向图用来为呼叫建模,其中用顶点表示每个电话号码,用有向边表示每次电话呼叫。表示呼叫的边是以发出呼叫的电话号码为起点,以接受呼叫的电话号码为终点。

一个小型的电话呼叫图如图 8-10a 所示,它表示 7 个电话号码。例如,这个图表示从 732-555-1234 到 732-555-9876 有 3 次呼叫和 2 次反向呼叫,但是从 732-555-4444 到其余 6 个号码(732-555-0011 除外)没有呼叫。当只关心是否有连接两个电话号码的呼叫时,就使用无向图,其中当两个号码之间有呼叫时,就用边连接这两个电话号码。这个版本的呼叫图如图 8-10b 所示。 ■

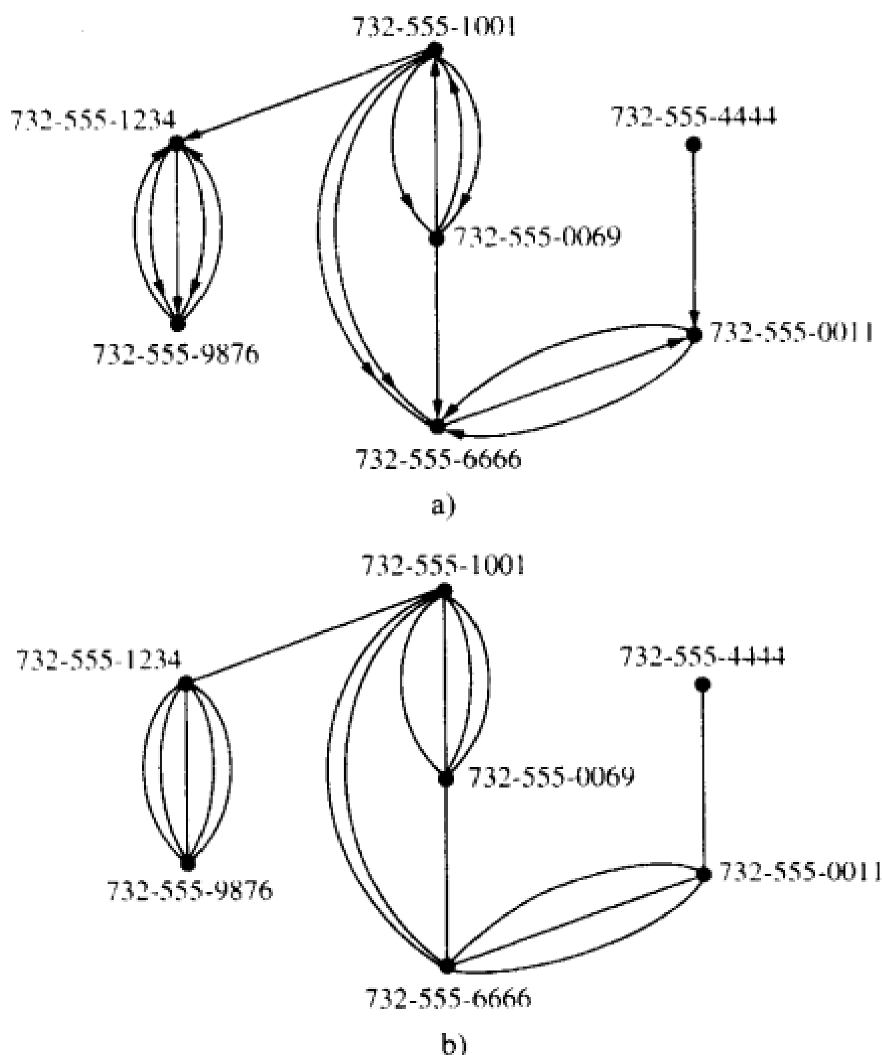


图 8-10 呼叫图

表示实际呼叫活动的呼叫图可以非常大。例如,在 AT&T 研究的一个呼叫图里,有大约 2 亿 9 千万个顶点和 40 亿条边,这个图表示在 20 天里进行的呼叫。在 8.4 节里将要进一步讨论呼叫图。

例 8 网络图 万维网可用有向图来建模,其中用顶点表示每个网页,并且若有从网页 a 指向网页 b 的链接,则有以 a 为起点以 b 为终点的边。因为几乎每秒钟都有新页面在网络上某处产生并且有其他页面被删除,所以网络图几乎是连续变化的。目前网络图有超过 10 亿个顶点和几百亿条边。许多人正在研究网络图的性质,以便更好地理解网络的特性。8.4 节将要继续讲解网络图,第 9 章将要解释网络爬虫(搜索引擎用它来产生网页的索引)是如何利用网络图的。 ■

例 9 优先图与并发处理 通过并发地执行某些语句,计算机程序可以执行得更快。重要的是避免语句执行时还要用到尚未执行语句的结果。语句与前面语句的相关性可以表示成有向图。用顶点表示每个语句,若在执行完第一个顶点所表示的语句之前不能执行第二个顶点所表示的

语句，则从第一个顶点到第二个顶点有一条边。这样的图称为优先图。图 8-11 显示计算机程序和优先图。例如，该图说明在执行语句 S_1、S_2 和 S_4 之前不能执行语句 S_5。■

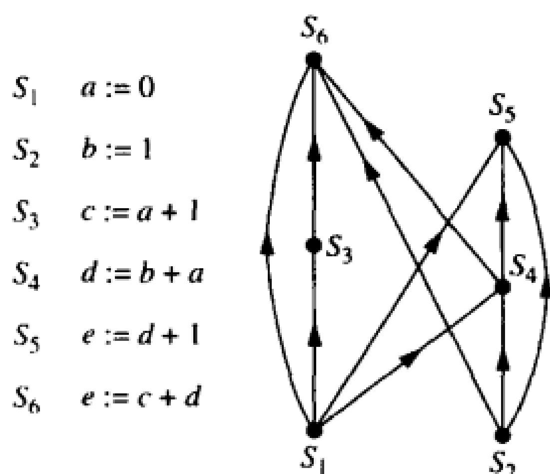
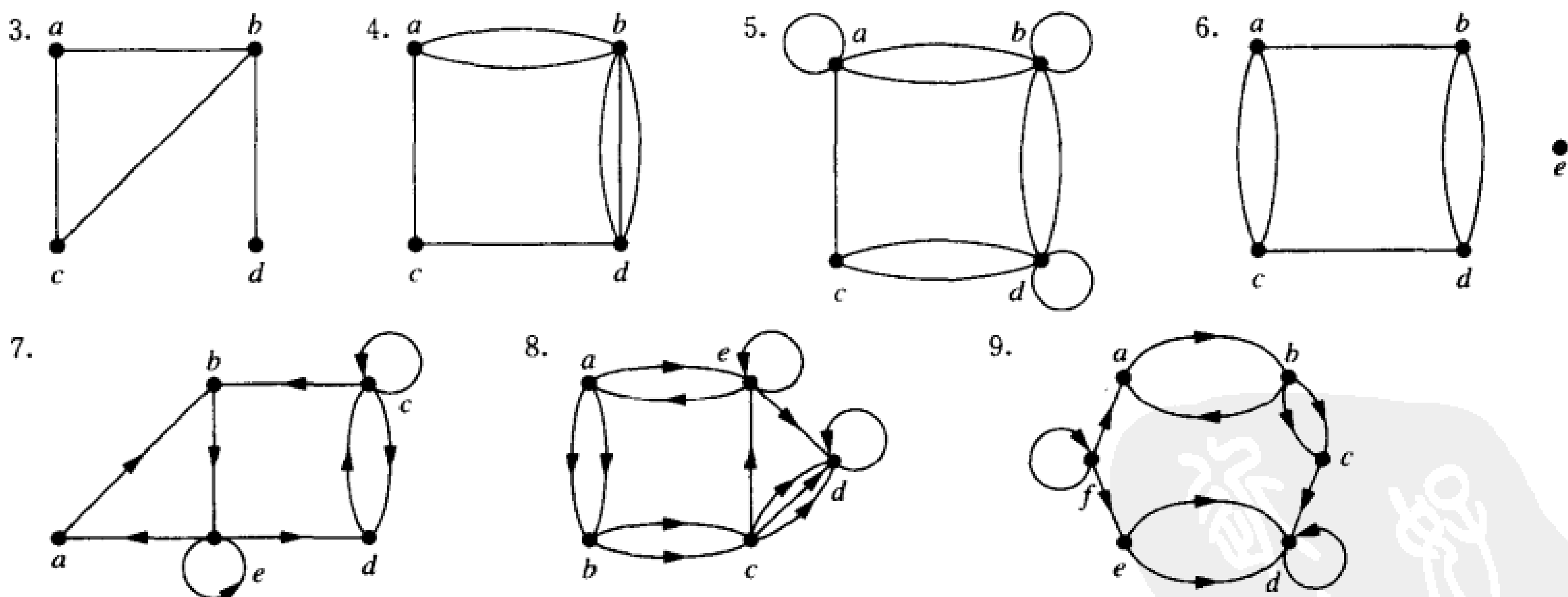


图 8-11 优先图

练习

1. 画出表示航空公司航线的图模型，并说出所用图的类型，其中每天有 4 个航班从波士顿到纽华克，2 个航班从纽华克到波士顿，3 个航班从纽华克到迈阿密，2 个航班从迈阿密到纽华克，1 个航班从纽华克到底特律，2 个航班从底特律到纽华克，3 个航班从纽华克到华盛顿，2 个航班从华盛顿到纽华克，1 个航班从华盛顿到迈阿密，其中
   a)若城市之间有航班(任何方向)，则在表示城市的顶点之间有边。
   b)对城市之间每个航班(任何方向)来说，在表示城市的顶点之间有边。
   c)对城市之间每个航班(任何方向)来说，在表示城市的顶点之间有边，以及有表示在迈阿密起飞和降落的特殊观光旅行的环。
   d)从表示航班出发城市的顶点到表示航班终止城市的顶点之间有边。
   e)对每个航班，从表示出发城市的顶点到表示终止城市的顶点之间有边。
2. 用什么类型的图来为大城市之间高速公路系统建模，其中
   a)若在城市之间有州际高速公路，则在表示城市的顶点之间有边。
   b)对城市之间每条高速公路，在表示城市的顶点之间有边。
   c)对城市之间每条高速公路，在表示城市的顶点之间有边；以及若有环城州际高速公路，则在表示该城的顶点上有环。

对练习 3～9，判断所示的图是否为简单图、多重图(非简单图)、伪图(非多重图)、有向图或有向多重图(非有向图)。



10. 对练习 3～9 里每个非简单无向图，找出使它变成简单图的可删除的边。
11. 集合 A_1，A_2，…，A_n 的交图对每个集合来说都有一个顶点，若两个集合有非空交集，则有一条边连接代表这两个集合的顶点。构造下列集合的交图。
    a) $A_1=\{0, 2, 4, 6, 8\}$，$A_2=\{0, 1, 2, 3, 4\}$，
    $A_3=\{1, 3, 5, 7, 9\}$，$A_4=\{5, 6, 7, 8, 9\}$，
    $A_5=\{0, 1, 8, 9\}$
    b) $A_1=\{\cdots, -4, -3, -2, -1, 0\}$，

$A_2=\{\cdots, -2, -1, 0, 1, 2, \cdots\}$,
$A_3=\{\cdots, -6, -4, -2, 0, 2, 4, 6, \cdots\}$,
$A_4=\{\cdots, -5, -3, -1, 1, 3, 5, \cdots\}$,
$A_5=\{\cdots, -6, -3, 0, 3, 6, \cdots\}$

c) $A_1=\{x \mid x<0\}$, $A_2=\{x \mid -1<x<0\}$,
$A_3=\{x \mid 0<x<1\}$,
$A_4=\{x \mid -1<x<1\}$, $A_5=\{x \mid x>-1\}$,
$A_6=\mathbf{R}$

12. 用图 8-6 里的栖息地重叠图来确定与鹰竞争的物种。
13. 构造 6 种鸟的栖息地重叠图，其中隐士鸫与旅鸫以及蓝松鸦竞争，旅鸫也与嘲鸫竞争，嘲鸫也与蓝松鸦竞争，以及鸸鸟与多毛啄木鸟竞争。
14. 画出熟人关系图表示 Tom 与 Patricia、Tom 与 Hope、Tom 与 Sandy、Tom 与 Amy、Tom 与 Marika、Jeff 与 Patricia、Jeff 与 Mary、Patricia 与 Hope、Amy 与 Hope 以及 Amy 与 Marika 互相认识，但是除上述以外，其他人与人之间互相不认识。
15. 可用图来表示两个人是否生活在同一时代。画出这样的图来表示本书前 4 章里有生平介绍的在 1900 年前去世的属于同一时代的数学家和计算机科学家。（如果在同一年里两个人都在世，就假设他们生活在同一时代。）
16. 在例 3 的影响图里谁影响 Fred？Fred 影响谁？
17. 构造公司董事会成员的影响图，主席影响技术总监、市场总监和营运总监；技术总监影响营运总监；市场总监影响营运总监；无人影响首席财务官或受其影响。
18. 在图 8-9 表示的循环赛里，队 4 击败哪些其他队？哪些队击败队 4？
19. 在循环赛里，老虎队击败蓝松鸦队，蓝松鸦队击败红衣主教队，红衣主教队击败北美金莺队，用有向图为这样的结果建模。
20. 画出电话号码 555-0011、555-1221、555-1333、555-8888、555-2222、555-0091 以及 555-1200 的呼叫图，假设从 555-0011 到 555-8888 有 3 次呼叫，从 555-8888 到 555-0011 有 2 次呼叫，从 555-2222 到 555-0091 有 2 次呼叫，从 555-1221 到其他每个号码有 2 次呼叫，从 555-1333 到 555-0011、555-1221 和 555-1200 各有 1 次呼叫。
21. 解释如何用一月份和二月份的电话呼叫图来确定改变了电话号码的人的新电话号码。
22. a)解释如何用图为网络里的电子邮件消息建模。
    b)描述为网络里具体某一周内电子邮件建模的图。
23. 如何用表示网络里发送电子邮件的图来找出最近改变了原来电子邮件地址的人？
24. 如何用表示网络里发送电子邮件的图来找出电子邮件的地址表，这样的地址表用来发送同样的消息给许多不同的电子邮件地址。
25. 描述一种表示婚姻的图模型。这个图有什么特殊性质？
26. 在例 9 的程序里，执行 S_6 之前必须执行哪些语句？（使用图 8-11 的优先图。）
27. 构造下列程序的优先图：

S_1：x：$=0$
S_2：x：$=x+1$
S_3：y：$=2$
S_4：z：$=y$
S_5：x：$=x+2$
S_6：y：$=x+z$
S_7：z：$=4$

28. 描述一种用来为航空公司的航线和航班时间建模的基于图的离散结构。（提示：给一个有向图添加

结构。）

29. 描述一种用来为群体里成对个人之间关系建模的基于图的离散结构，其中每个人可能喜欢或者不喜欢另一人，或者中立，而反过来的关系可以是不同的。（提示：给一个有向图添加结构。分别处理表示两个人的顶点之间的反向边。）

8.2 图的术语

8.2.1 引言

本节将介绍图论的一些基本词汇。当解决许多不同类型的问题时可以使用这些词汇。其中一个这样的问题涉及判定能否把图画在平面里，使得没有两条边是交叉的。另一个例子是判定两个图是否具有顶点之间的一一对应，使得这样的对应能够产生边之间的一一对应。我们还将介绍在例子和模型里经常用到的几种重要的图族。

8.2.2 基本术语

首先给出描述无向图的顶点和边的一些术语。

定义 1 若 $\{u, v\}$ 是无向图 G 的边，则两个顶点 u 和 v 称为在 G 里邻接（或相邻）。若 $e=\{u, v\}$，则边 e 称为关联顶点 u 和 v。也可以说边 e 连接 u 和 v。顶点 u 和 v 称为边 $\{u, v\}$ 的端点。

为了反映出有多少条边关联着一个顶点，有下述的定义。

定义 2 在无向图里顶点的度是与该顶点关联的边的数目，例外的情形是，顶点上的环为顶点的度做出双倍贡献。顶点 v 的度表示成 $\deg(v)$。

例 1 如图 8-12 所示，图 G 和图 H 的顶点的度是什么？

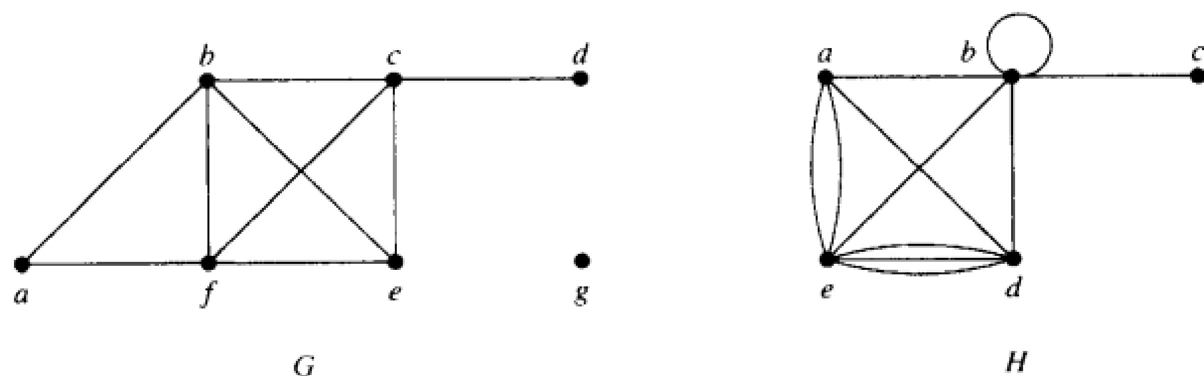


图 8-12 无向图 G 和 H

解 在 G 里，$\deg(a)=2$，$\deg(b)=\deg(c)=\deg(f)=4$，$\deg(d)=1$，$\deg(e)=3$，$\deg(g)=0$。在 H 里，$\deg(a)=4$，$\deg(b)=\deg(e)=6$，$\deg(c)=1$，$\deg(d)=5$。 ■

把度为 0 的顶点称为孤立的。因此孤立点不与任何顶点相邻。例 1 里图 G 的顶点 g 是孤立的。顶点是悬挂的，当且仅当它的度是 1。因此悬挂点恰与 1 个其他顶点相邻。例 1 里图 G 的顶点 d 是悬挂的。

当对图 $G=(V, E)$ 的所有顶点的度求和时，得出了什么？每条边都为顶点的度之和贡献 2，因为一条边恰好关联 2 个（可能相同）顶点。这意味着顶点的度之和是边数的 2 倍。下面描述握手定理，这是因为在一条边上有两个端点可以类比为一次握手涉及两只手这种情形。

定理 1 握手定理 设 $G=(V, E)$ 是有 e 条边的无向图，则

$$2e = \sum_{v\in V}\deg(v)$$

（注意即使出现多重边和环，这个式子也仍然成立。）

例 2 一个具有 10 个顶点而且每个顶点的度都为 6 的图，有多少条边？

解 因为顶点的度之和是 $6 \cdot 10=60$，所以 $2e=60$。因此 $e=30$。■

定理 1 说明无向图顶点的度之和是偶数。这可以推导出许多结论，其中一个结论作为定理 2 给出。

定理 2 无向图有偶数个奇数度顶点。

证 在无向图 $G=(V,E)$ 里，设 V_1 和 V_2 分别是偶数度顶点和奇数度顶点的集合。于是

$$2e=\sum_{v\in V}\deg(v)=\sum_{v\in V_1}\deg(v)+\sum_{v\in V_2}\deg(v)$$

因为对 $v\in V_1$ 来说，$\deg(v)$ 是偶数，所以上面等式右端的第一项是偶数。另外，上面等式右端的两项之和是偶数，因为和是 $2e$。因此，和里的第二项也是偶数。因为在第二项的和里，所有的项都是奇数，所以必然有偶数个这样的项。因此，有偶数个奇数度顶点。□

对带有有向边的图来说也有一些有用的术语。

定义 3 当 (u,v) 是带有有向边的图 G 的边时，说 u *邻接到* v，而且说 v *从* u *邻接*。顶点 u 称为 (u,v) 的*起点*，v 称为 (u,v) 的*终点*。环的起点和终点是相同的。

因为带有有向边的图的边是有序对，所以这时顶点度的定义细化成反映这个顶点作为起点和作为终点的不同度数。

定义 4 在带有向边的图里，顶点 v 的*入度*（表示成 $\deg^-(v)$）是以 v 作为终点的边数。顶点 v 的*出度*（表示成 $\deg^+(v)$）是以 v 作为起点的边数（注意顶点上的环对这个顶点的入度和出度的贡献都是 1）。

例 3 求出图 8-13 所示带有向边的图 G 里每个顶点的入度和出度。

解 G 里的入度是：$\deg^-(a)=2$，$\deg^-(b)=2$，$\deg^-(c)=3$，$\deg^-(d)=2$，$\deg^-(e)=3$，$\deg^-(f)=0$。出度是：$\deg^+(a)=4$，$\deg^+(b)=1$，$\deg^+(c)=2$，$\deg^+(d)=2$，$\deg^+(e)=3$，$\deg^+(f)=0$。■

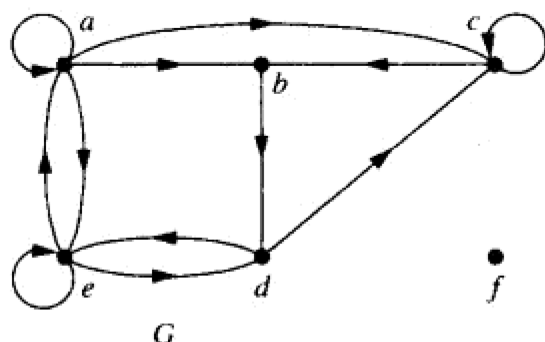


图 8-13 有向图 G

因为每条边都有一个起点和一个终点，所以在带有向边的图里，所有顶点的入度之和等于出度之和。这两个和都等于图的边数。把这个结果表述成定理 3。

定理 3 设 $G=(V,E)$ 是带有向边的图。于是

$$\sum_{v\in V}\deg^-(v)=\sum_{v\in V}\deg^+(v)=|E|$$

带有向边的图有许多性质是不依赖于边的方向的。因此忽略这些方向经常是有用处的。忽略边的方向后得出的无向图称为*底无向图*。带有向边的图与它的底无向图有相同的边数。

8.2.3 一些特殊的简单图

下面要介绍几类简单图。这些图常常用作例子，来自于许多应用。

例 4 **完全图** n 个顶点的*完全图*（表示成 K_n）是在每对不同顶点之间都恰有一条边的简单图。对 $n=1, 2, 3, 4, 5, 6$ 来说，图 8-14 显示图 K_n。■

例 5 **圈图** *圈图* $C_n(n\geqslant 3)$ 是由 n 个顶点 $v_1, v_2, \cdots, v_n$ 以及边 $\{v_1, v_2\}, \{v_2, v_3\}, \cdots, \{v_{n-1}, v_n\}, \{v_n, v_1\}$ 组成的。图 8-15 显示圈图 C_3, C_4, C_5, C_6。■

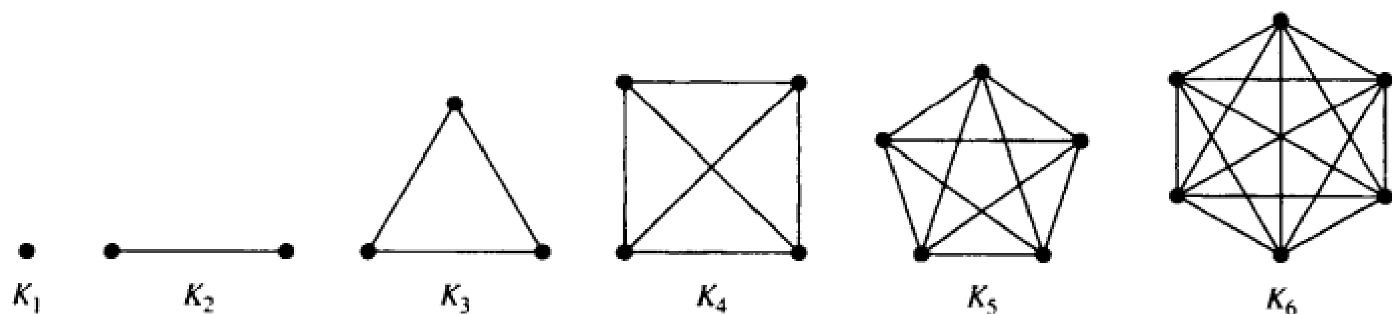


图 8-14 图 K_n，$n=1$，2，3，4，5，6

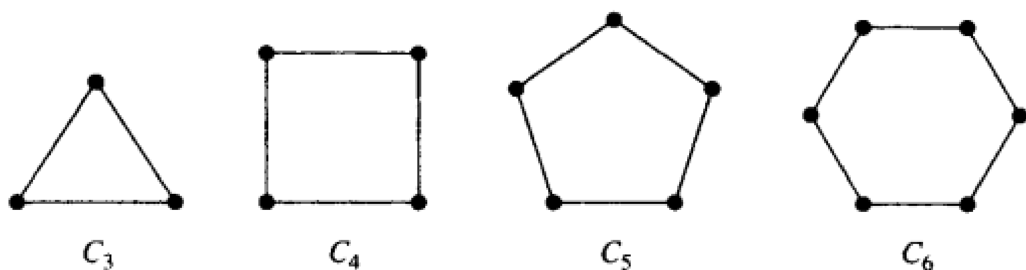


图 8-15 圈图 C_3，C_4，C_5，C_6

例 6 轮图 对 $n\geqslant 3$ 来说，当给圈图 C_n 添加另一个顶点，而且把这个新顶点与 C_n 里 n 个顶点逐个连接时，就得出轮图 W_n。图 8-16 显示轮图 W_3，W_4，W_5，W_6。 ■

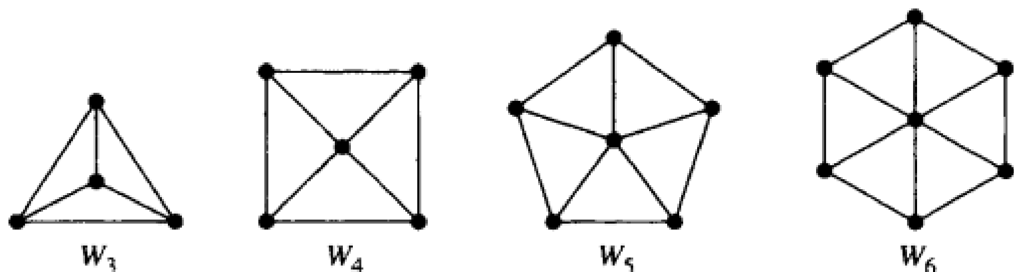


图 8-16 轮图 W_3，W_4，W_5，W_6

例 7 n 立方体图 n 立方体图(表示成 Q_n)是用顶点表示 2^n 个长度为 n 的位串的图。两个顶点相邻，当且仅当它们所表示的位串恰恰相差一位。图 8-17 显示图 Q_1、Q_2、Q_3。注意可以从 n 立方体图 Q_n 来构造($n+1$)立方体图 Q_{n+1}，方法是建立 Q_n 的两个副本，在 Q_n 的一个副本的顶点标记前加 0，在 Q_n 的另一个副本的顶点标记前加 1，并且加入连接那些标记只在第一位不同的两个顶点的边。在图 8-17 中，从 Q_2 来构造 Q_3，方法是画出 Q_2 的两个副本作为 Q_3 的顶面和底面，在底面每个顶点的标记前加 0，在顶面每个顶点的标记前加 1。 ■

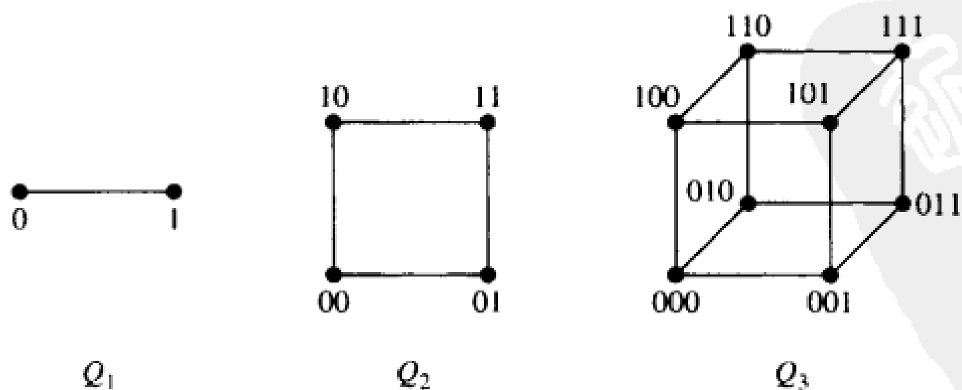


图 8-17 对于 $n=1$，2，3 的 n 立方体图 Q_n

8.2.4 偶图

有时可以把图的顶点分成两个不相交的子集，使得每条边都连接一个子集里的顶点与另一个子集里的顶点。例如，考虑一下表示村民之间的婚姻关系的图，其中用顶点表示每个人，用边表示婚姻。在这个图里，每条边都连接表示男人的顶点子集里的顶点与表示女人的顶点子集里的顶点。这引出了定义 5。

定义 5 若把简单图 G 的顶点集分成两个不相交的非空集合 V_1 和 V_2，使得图里的每一条边都连接着 V_1 里的一个顶点与 V_2 里的一个顶点(因此 G 里没有边是连接着 V_1 里的两个顶点或 V_2 里的两个顶点)，则 G 称为偶图[⊖]。

在例 8 里说明 C_6 是偶图，在例 9 里说明 K_3 不是偶图。

例 8 图 8-18 所示的 C_6 是偶图，因为把它的顶点集分成两个集合 $V_1=\{v_1, v_3, v_5\}$ 和 $V_2=\{v_2, v_4, v_6\}$，C_6 的每一条边都连接 V_1 里的一个顶点与 V_2 里的一个顶点。 ■

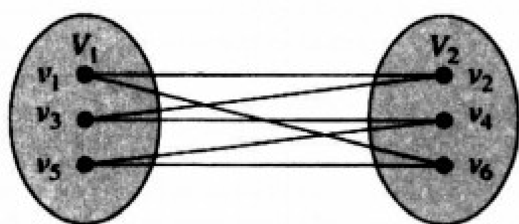


图 8-18 C_6 是偶图

例 9 K_3 不是偶图。为了看明白这一点，注意若把 K_3 的顶点集分成两个不相交的集合，则两个集合之一必然包含两个顶点。假如这个图是偶图，那么这两个顶点就不能用边连接，但是在 K_3 里每一个顶点都用边连接到其他每个顶点。 ■

例 10 图 8-19 所示的图 G 和 H 是否为偶图?

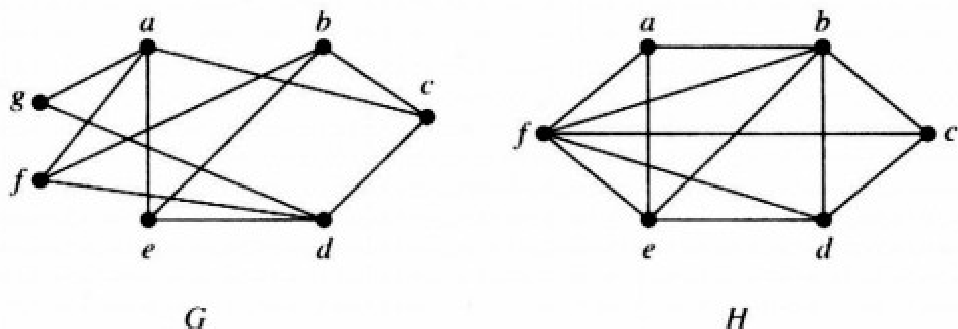


图 8-19 无向图 G 和 H

解 图 G 是偶图，因为它的顶点集是两个不相交集合 $\{a, b, d\}$ 和 $\{c, e, f, g\}$ 的并，每条边都连接一个子集里的一个顶点与另一个子集里的一个顶点。(注意对偶图 G 来说，不必让 $\{a, b, d\}$ 里每一个顶点与 $\{c, e, f, g\}$ 里每一个顶点都相邻。例如 b 与 g 就不相邻。)

图 H 不是偶图，因为它的顶点集不能分成两个子集，使得边都不连接同一个子集的两个顶点(读者应当通过考虑顶点 a、b、f 来验证它)。 ■

一个图是偶图，当且仅当可以用至多两种颜色来为这个图的顶点着色，使得任何两个相邻的顶点都不同色。将在 8.8 节里研究图的着色。另外，一个图是偶图，当且仅当不能经过奇数条不同的边来从一个顶点出发又返回这个顶点。当在 8.4 节里讨论通路与回路时，将更精确地说明这个概念。

例 11 完全偶图 完全偶图 $K_{m,n}$ 是顶点集分成分别含有 m 和 n 个顶点的两个子集的图。两个顶点之间有边当且仅当一个顶点属于第一个子集而另一个顶点属于第二个子集。图 8-20 显示完全偶图 $K_{2,3}$，$K_{3,3}$，$K_{3,5}$ 以及 $K_{2,6}$。 ■

⊖ 也称为二部图、二分图。——译者注

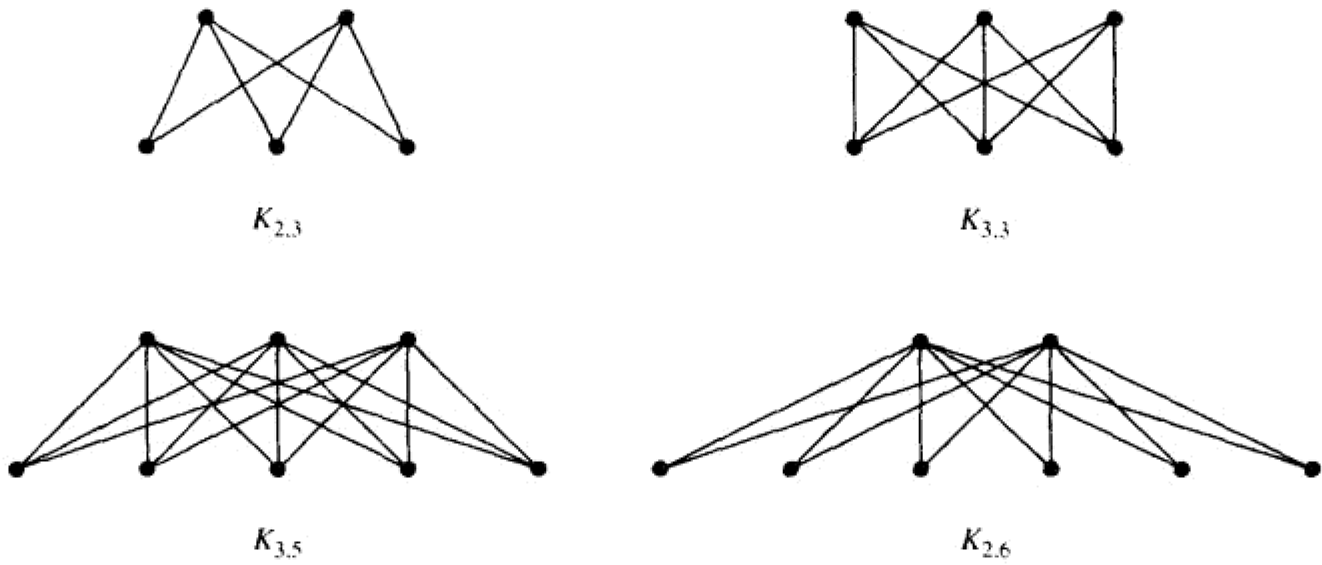


图 8-20 部分完全偶图

8.2.5 特殊类型的图的一些应用

以下将要说明如何把特殊类型的图用在数据通信和并行处理的模型里。

例 12 局域网 在一座大楼里，像小型计算机和个人电脑这样的各种计算机，以及像打印机和绘图仪这样的外设，都可以用*局域网*来连接。一些这样的网络是基于*星形拓扑*，其中所有设备都连接到中央控制设备。局域网用图 8-21a 所示的完全偶图 $K_{1,n}$ 来表示。通过中央控制设备在设备间传输信息。

其他局域网是基于*环形拓扑*，其中每个设备恰恰连接到两个其他设备。带环形拓扑的局域网用图 8-21b 所示的 n 圈图 C_n 来建模。消息围绕着圈从设备送到设备，直到抵达消息目的地为止。

最后，一些局域网采用这两种拓扑的混合形式。消息是围绕着环或通过中央设备来传送。这样的冗余使得网络更加可靠。带冗余的局域网用图 8-21c 所示的轮图 W_n 来建模。 ■

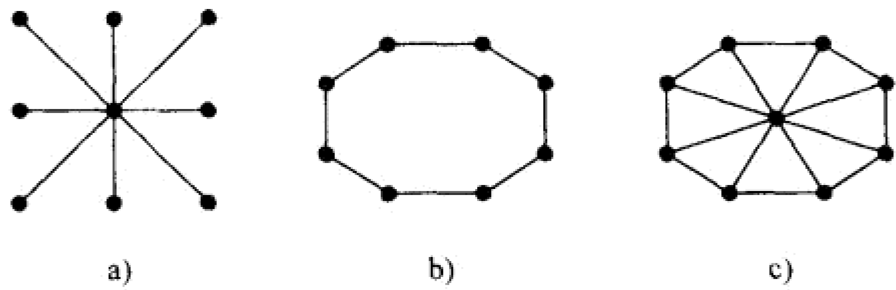


图 8 21 局域网的星形、环形以及混合拓扑

例 13 并行计算的互联网络 直到最近，计算机执行程序还是一次完成一个操作。因此，为解决问题而写的算法都设计成一次执行一步；这样的算法称为*串行的*(几乎所有本书描述的算法都是串行的)。不过，像气象模拟、医学图像分析以及密码分析这样的许多高强度计算问题，即使在超级计算机上，也不能通过串行操作在合理的时间范围内解决。另外，计算机执行基本操作的速度，还存在物理的限制，所以总是有问题不能用串行操作在合理的时间范围内解决。

*并行处理*利用包含许多自带内存的单处理器的计算机来帮助克服串行计算机的局限性。*并行算法*把问题分成可并发解决的若干子问题，于是可以设计并行算法，用带有多处理器的计算机来快速解决问题。在并行算法里，单个指令流控制着算法的执行，包括把子问题送到不同的处理器，以及把子问题的输入和输出定向到适当的处理器。

采用并行处理时，一个处理器需要另一个处理器产生的输出。因此处理器需要互联。可用适当类型的图来表示带有多重处理器的计算机里处理器的互联网络。在以下讨论里将要描述最常用类型的并行处理器互联网络。用来实现具体并行算法的互联网络的类型，差异在于处理器之间交换数据的需求量和所需要的速度，当然还有可用的硬件等。

最简单却又最昂贵的网络互联处理器，包含每对处理器之间的双向连接。当有 n 个处理器时，这样的网络表示成 n 个顶点上的完全图 K_n。不过，这种类型的互联网络有严重的问题，它所需要的连接数太大。实际上，处理器的直接连接数目是有限的，所以当处理器数很大时，处理器不能直接连接到所有其他处理器。例如，当有 64 个处理器时，就需要 $C(64, 2)=2016$ 个连接，每个处理器都得直接连接到其他 63 个处理器。

另一方面，互联 n 个处理器的最简单方式或许是使用称为*线性阵列*的排列方式。除 P_1 和 P_n 外的每个处理器 P_i 都通过双向连接来连接到相邻处理器 P_{i-1} 和 P_{i+1}。P_1 只连接 P_2，P_n 只连接 P_{n-1}。图 8-22 显示 6 处理器线性阵列。这种方式的缺点是为了让处理器共享信息，有时需要使用大量的称为 **hop** 的中间连接。

栅格网络(或*二维阵列*)是通用的互联网络。在这样的网络里，处理器个数是一个完全平方数，比方说 $n=m^2$。n 个处理器标记成 $P(i, j)$，$0\leqslant i\leqslant m-1$，$0\leqslant j\leqslant m-1$。双向连接把处理器 $P(i, j)$连接到它的 4 个相邻处理器 $P(i\pm1, j)$和 $P(i, j\pm1)$，只要这些处理器是在栅格里。(注意，栅格角上的 4 个处理器只有 2 个相邻处理器，边界上其他处理器只有 3 个相邻处理器；有时也用每个处理器恰有 4 个连接的变种的栅格网络；见本节末的练习 52。)栅格网络限制每个处理器的连接数。某些成对处理器之间的通信需要 $O(\sqrt{n})=O(m)$个中间连接(见本节末的练习 53)。图 8-23 表示 16 个处理器的栅格网络的图。



图 8-22　6 个处理器的线性阵列

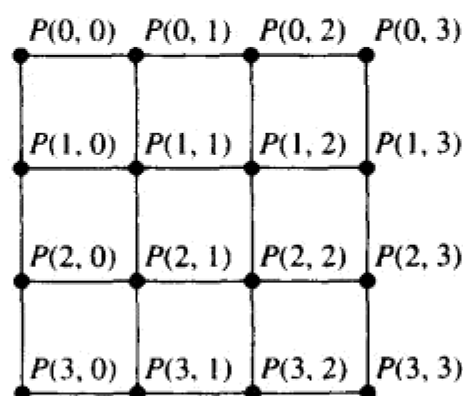


图 8-23　16 个处理器的栅格网络

*超立方体*是互联网络的一个重要类型。对这样的网络来说，处理器数是 2 的幂，$n=2^m$。n 个处理器标记成 P_0，P_1，…，P_{n-1}。每个处理器都有到 m 个其他处理器的双向连接。连接到处理器 P_i 上的处理器，其下标的二进制表示与 i 的二进制表示恰恰相差 1 位。超立方体网络在每个处理器的直接连接数与保证处理器通信的中间连接数之间取得了平衡。已经用超立方体网络建造了许多计算机，而且用超立方体网络设计了许多算法。n 立方体图 Q_n 表示带 n 个处理器的超立方体网络。图 8-24 显示 8 个处理器的超立方体网络(图 8-24 中画 Q_3 的方式与图 8-17 所示的方式不同)。 ■

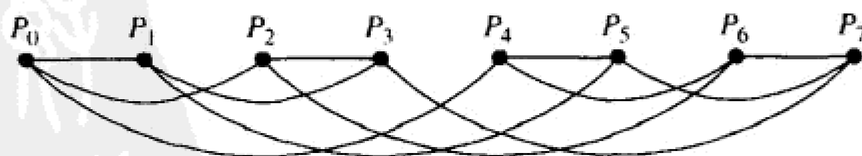


图 8-24　8 个处理器的超立方体网络

8.2.6 从旧图到新图

有时解决问题只需要图的一部分。例如，只关心大型计算机网络中涉及纽约、丹佛、底特律以及亚特兰大的计算机中心的那一部分。其他的计算机中心以及不连接到这 4 个具体计算机中心里任何 2 个的所有电话线可以忽略。在大网络的图模型里，可以删除除这 4 处外的计算机中心所对应的顶点，可以删除所有与所删除顶点关联的边。删除后所剩下的图称为原图的子图。

定义 6 图 $G=(V, E)$ 的子图是图 $H=(W, F)$，其中 $W\subseteq V$ 而且 $F\subseteq E$。

例 14 图 8-25 所示的图 G 是 K_5 的子图。 ■

可以用各种方式组合两个或更多的图。把这些图的所有顶点和边所组成的图称为这些图的并图。两个简单图的并图的定义如下。

定义 7 两个简单图 $G_1=(V_1, E_1)$ 和 $G_2=(V_2, E_2)$ 的并图是带有顶点集 $V_1\cup V_2$ 和边集 $E_1\cup E_2$ 的简单图。G_1 和 G_2 的并图表示成 $G_1\cup G_2$。

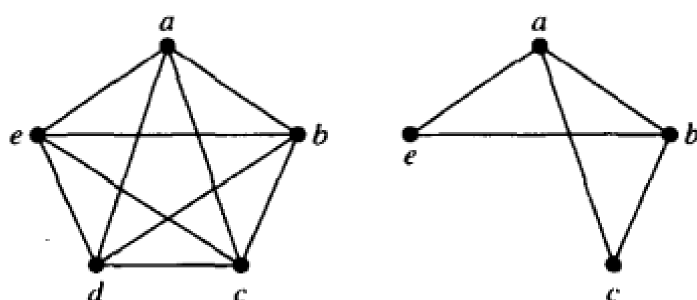


图 8-25 K_5 的子图

例 15 求图 8-26a 所示的图 G_1 和 G_2 的并图。

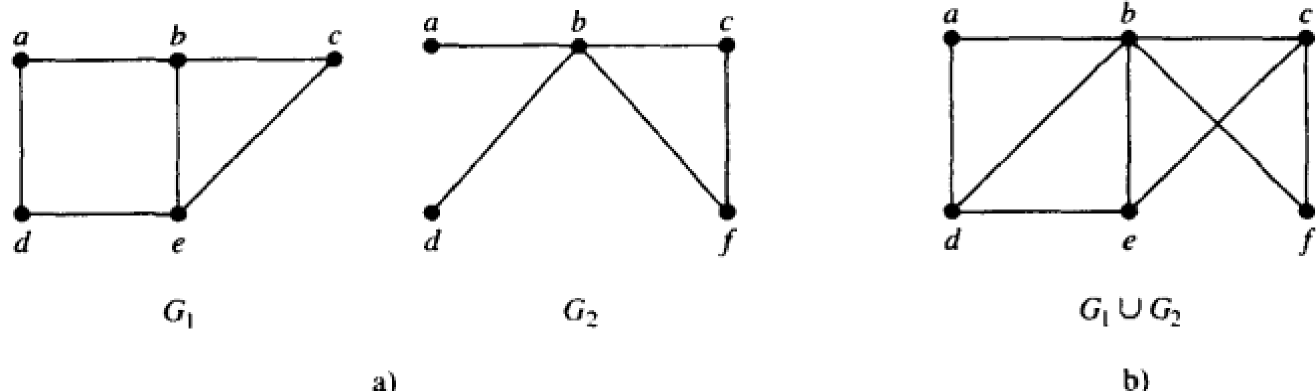
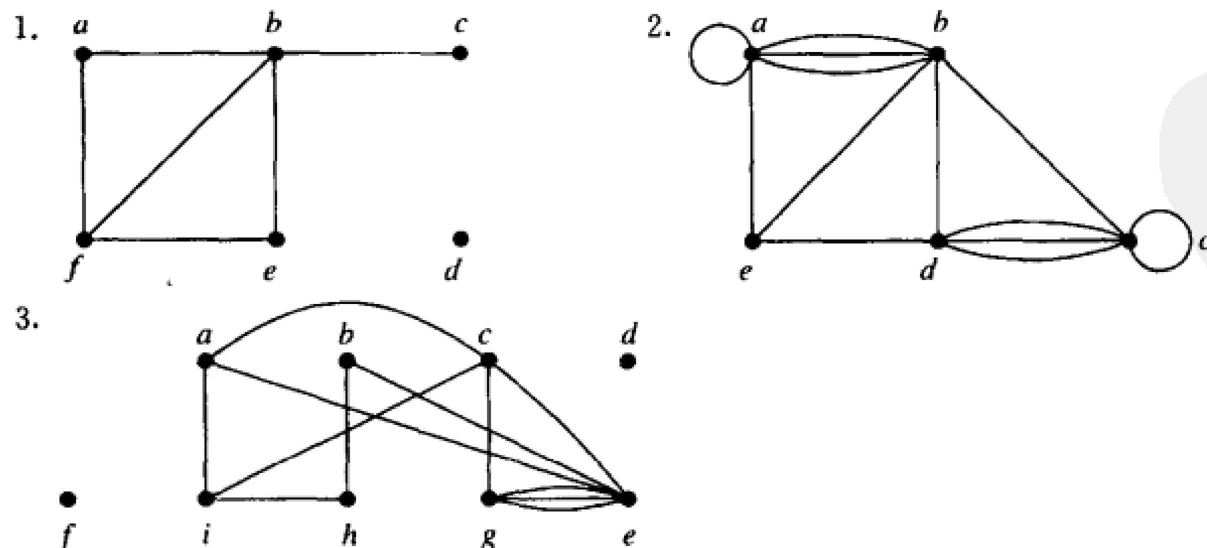


图 8-26 a)简单图 G_1 和 G_2；b)它们的并 $G_1\cup G_2$

解 并图 $G_1\cup G_2$ 的顶点集是两个顶点集的并，即$\{a, b, c, d, e, f\}$。并图的边集是两个边集的并。并图显示在图 8-26b 中。 ■

练习

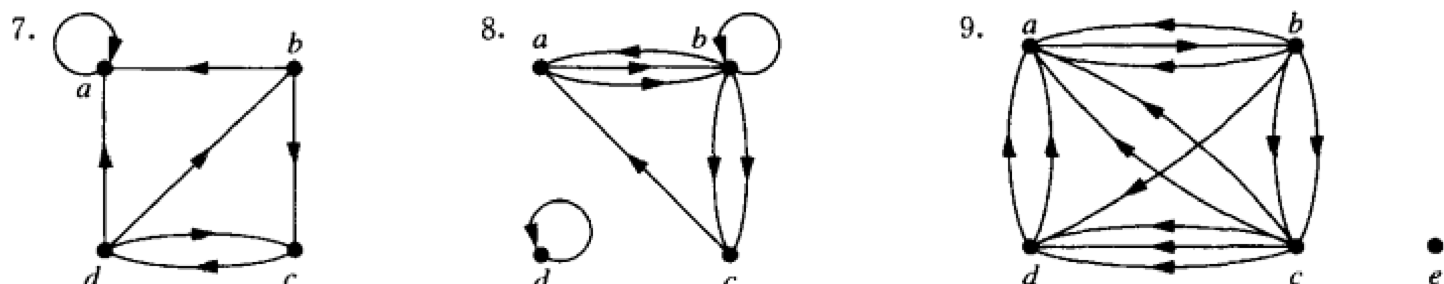
在练习 1～3 里，求所给无向图的顶点数、边数以及每个顶点的度。指出所有孤立点和悬挂点。



4. 求练习 1～3 里每个图的顶点度之和，并验证它等于图里边数的 2 倍。
5. 带有 15 个顶点而且每个顶点的度都为 5 的简单图能否存在？

6. 证明：在一次聚会上全体人员的握手次数之和是偶数。假设无人自己与自己握手。

在练习 7～9 里，对给定的有向多重图来说，确定顶点数和边数，并求出每个顶点的入度和出度。

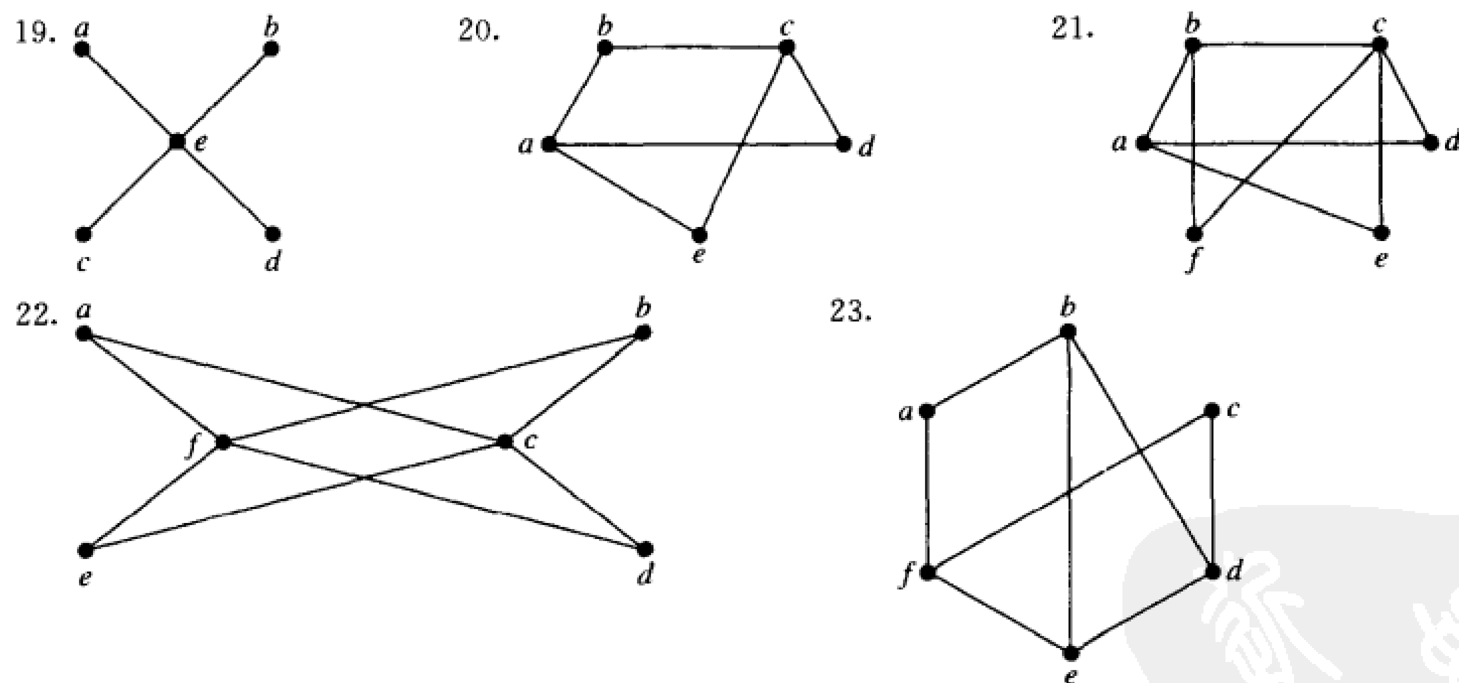


10. 对练习 7～9 里的每个图来说，直接确定每个顶点的入度之和与出度之和。证明：它们都等于图里的边数。
11. 构造出图 8-13 里带有向边的图的底图。
12. 在熟人关系图里(其中顶点表示世界上所有的人)，顶点的度表示什么？在这个图里孤立点和悬挂点表示什么？在一项研究中估计在这个图里顶点的平均度是 1000。就这个模型而言，这意味着什么？
13. 在合作图里，顶点的度表示什么？孤立点和悬挂点表示什么？
14. 在好莱坞图里，顶点的度表示什么？孤立点和悬挂点表示什么？
15. 在 8.1 节例 7 所描述的电话呼叫图里，顶点的入度和出度表示什么？在这个图的无向版本里，顶点的度表示什么？
16. 在 8.1 节例 8 所描述的网络图里，顶点的入度和出度表示什么？
17. 在为循环赛建模的有向图里，顶点的入度和出度表示什么？
18. 画出下列图。

a) K_7　　b) $K_{1,8}$　　c) $K_{4,4}$

d) C_7　　e) W_7　　f) Q_4

在练习 19～23 里，确定图是否为偶图。



24. 对哪些 n 值来说，下列图是偶图？

a) K_n　　b) C_n　　c) W_n　　d) Q_n

25. 下列图有多少个顶点和多少条边？

a) K_n　　b) C_n　　c) W_n

d) $K_{m,n}$　　e) Q_n

26. 若图的顶点有度 4，3，3，2，2，则它有多少条边？画出这样的图。
27. 若图的顶点有度 5，2，2，2，2，1，则它有多少条边？画出这样的图。
28. 是否存在带有下列度的有 6 个顶点的简单图？若有，画出这样的图。

a)0，1，2，3，4，5　　b)1，2，3，4，5，6

c)2，2，2，2，2，2　　d)3，2，3，2，3，2

e)3，2，2，2，2，3　　f)1，1，1，1，1，1

g)3，3，3，3，3，5　　h)1，2，3，4，5，5

29. 是否存在带有下列度的有 5 个顶点的简单图？若有，画出这样的图。

a)3，3，3，3，2　　b)1，2，3，4，5

c)1，2，3，4，4　　d)3，4，3，4，3

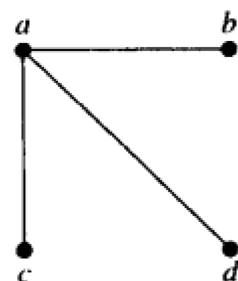
e)0，1，2，2，3　　f)1，1，1，1，1

30. 至少带有 1 个顶点的 K_2 的子图有多少个？

31. 至少带有 1 个顶点的 K_3 的子图有多少个？

32. 至少带有 1 个顶点的 W_3 的子图有多少个？

33. 画出右面这个图的所有子图。



34. 设 G 是带有 v 个顶点和 e 条边的图。设 M 是 G 的顶点的最大度，设 m 是 G 的顶点的最小度。证明：

a)$2e/v \geqslant m$　　b)$2e/v \leqslant M$

若简单图每个顶点的度都相等，则这个图称为正则的。若正则图每个顶点的度都为 n，则这个图称为 n 正则的。

35. 对哪些 n 值来说，下列图是正则图？

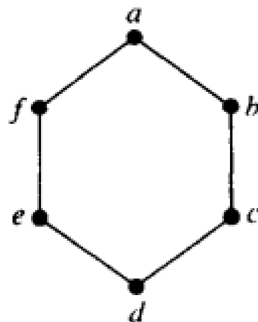
a)K_n　　b)C_n　　c)W_n　　d)Q_n

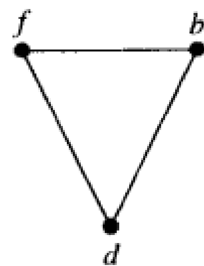
36. 对哪些 m 和 n 的值来说，$K_{m,n}$ 是正则图？

37. 度都为 4 而且带有 10 条边的正则图有多少个顶点？

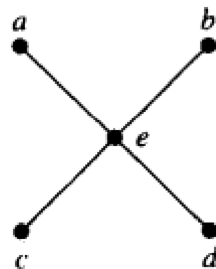
在练习 38～40 里，求给定的简单图对的并图（假设带有相同端点的边是相同的）。

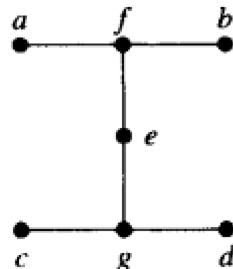
38.



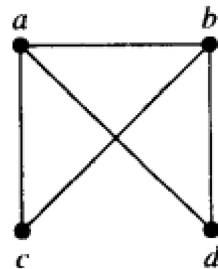


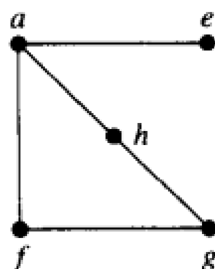
39.





40.





41. 简单图 G 的补图 $\overline{G}$ 与 G 有相同的顶点。两个顶点在 $\overline{G}$ 里相邻，当且仅当它们在 G 里不相邻。求下列图。

a)$\overline{K_n}$　　b)$\overline{K_{m,n}}$　　c)$\overline{C_n}$　　d)$\overline{Q_n}$

42. 若 G 是有 15 条边的简单图而且 $\overline{G}$ 有 13 条边，则 G 有多少个顶点？

43. 若简单图 G 有 v 个顶点和 e 条边，则 $\overline{G}$ 有多少条边？

*44. 证明：若 G 是有 v 个顶点和 e 条边的偶简单图，则 $e \leqslant v^2/4$。

45. 证明：若 G 是有 n 个顶点的简单图，则 G 和 $\overline{G}$ 的并图是 K_n。

*46. 描述判定图是否为偶图的算法。

表示成 G^c 的有向图 $G=(V, E)$ 的逆图是图 (V, F)，其中 $(u, v) \in F$，当且仅当 $(v, u) \in E$。

47. 描述 8.1 节练习 7～9 里每个图的逆图。

48. 证明：每当 G 是有向图时，就有 $(G^c)^c=G$。
49. 证明：图 G 是它自身的逆图，当且仅当 G 所关联的关系(参见 7.3 节)是对称的。
50. 把有向图的逆图的定义推广到有向多重图的逆图的概念。
51. 画出互联 9 个并行处理器的栅格网络。
52. 在互联 $n=m^2$ 个处理器的栅格网络的变种里，处理器 $P(i, j)$ 连接 4 个处理器 $P((i\pm 1) \bmod m, j)$，$P(i, (j\pm 1) \bmod m)$，使得连接沿栅格的边卷绕。画出有 16 个处理器的这种变种的栅格网络。
53. 证明：在 $n=m^2$ 个处理器的栅格网络里，用 $O(\sqrt{n})=O(m)$ 个 hop 就能让每一对处理器互相通信。

8.3 图的表示和图的同构

8.3.1 引言

图的表示方式有很多种。在本章里将看到选择最方便的表示会有助于对图的处理。本节将要说明如何用多种不同的方式来表示图。

有时两个图在这样一种意义下恰好具有相同的形式，就是在两个图的顶点之间存在着一一对应，这个对应保持边的对应关系。在这种情形下就说两个图是*同构的*。确定两个图是否同构，这是本节里将要研究的重要图论问题。

8.3.2 图的表示

表示不带多重边的图的一种方式是列出这个图的所有边。另一种表示不带多重边的图的方式是用*邻接表*，它规定与图的每个顶点相邻的顶点。

例 1 用邻接表描述图 8-27 所示的简单图。

解 表 8-2 列出与图的每个顶点相邻的顶点。 ■

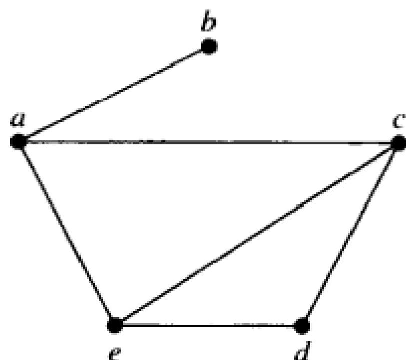


图 8-27 简单图

表 8-2 简单图的边表

顶点	相邻顶点
a	b, c, e
b	a
c	a, d, e
d	c, e
e	a, c, d

例 2 通过列举在图的每个顶点上发出的边的终点，表示图 8-28 所示的有向图。

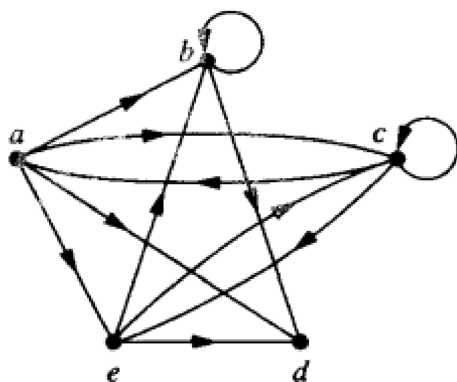


图 8-28 一个有向图

表 8-3 有向图的边表

起点	终点
a	b, c, d, e
b	b, d
c	a, c, e
d	
e	b, c, d

解 表 8-3 表示图 8-28 所示的有向图。 ■

8.3.3 邻接矩阵

若图里有许多边，则把图表示成边表或邻接表，会不便于执行图的算法。为了简化计算，可用矩阵表示图。在此将给出两种类型的常用的表示图的矩阵。一种类型是基于顶点的相邻关系，另一种类型是基于顶点与边的关联关系。

假设 $G=(V, E)$ 是简单图，其中 $|V|=n$。假设把 G 的顶点任意地排列成 v_1，v_2，…，v_n。对这个顶点表来说，G 的邻接矩阵 $\mathbf{A}$(或 $\mathbf{A}_G$)是一个 $n\times n$ 的 0－1 矩阵，它满足这样的性质：当 v_i 和 v_j 相邻时第(i, j)项是 1，当 v_i 和 v_j 不相邻时第(i, j)项是 0。换句话说，若邻接矩阵是 $\mathbf{A}=[a_{ij}]$，则

$$a_{ij}=\begin{cases}1, & 若\{v_i,v_j\}是G的一条边\\ 0, & 否则\end{cases}$$

注意，图的邻接矩阵依赖于所选择的顶点的顺序。因此带 n 个顶点的图有 $n!$ 个不同的邻接矩阵，因为 n 个顶点有 $n!$ 个不同的顺序。

简单图的邻接矩阵是对称的，即 $a_{ij}=a_{ji}$，因为当 v_i 和 v_j 相邻时，这两个项都是 1，否则都是 0。另外，因为简单图无环，所以每一项 a_{ii} 都是 0，$i=1, 2, 3, \cdots, n$。

注意，当图里的边相对少时，邻接矩阵是*稀疏矩阵*，即只有很少的非 0 项的矩阵。可以用特殊的方法来表示和计算这样的矩阵。另外，当表示和处理这样的图时，用边表有时是更有效的。

例 3 用邻接矩阵表示图 8-29 所示的图。

解 把顶点排列成 a，b，c，d。表示这个图的矩阵是

$$\begin{bmatrix}0&1&1&1\\1&0&1&0\\1&1&0&0\\1&0&0&0\end{bmatrix}$$ ■

例 4 画出带有相对于顶点顺序 a，b，c，d 的邻接矩阵的图。

$$\begin{bmatrix}0&1&1&0\\1&0&0&1\\1&0&0&1\\0&1&1&0\end{bmatrix}$$

解 图 8-30 显示带有这个邻接矩阵的图。 ■

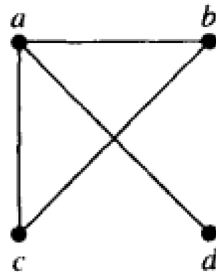


图 8-29 简单图

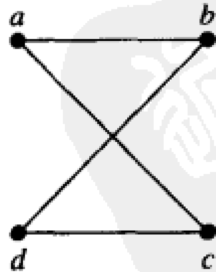


图 8-30 带有给定的邻接矩阵的图

邻接矩阵也可用来表示带环和多重边的无向图。把顶点 a_i 上的环表示成邻接矩阵第(i, i)位置上的 1。当出现多重边时，邻接矩阵不再是 0－1 矩阵，这是因为邻接矩阵的第(i, j)项等于与$\{a_i, a_j\}$关联的边数。包括多重图与伪图在内的所有无向图都具有对称的邻接矩阵。

例 5 用邻接矩阵表示图 8-31 所示的伪图。

解 顶点顺序为 a，b，c，d 的邻接矩阵是

$$\begin{bmatrix} 0 & 3 & 0 & 2 \\ 3 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 2 & 1 & 2 & 0 \end{bmatrix}$$

■

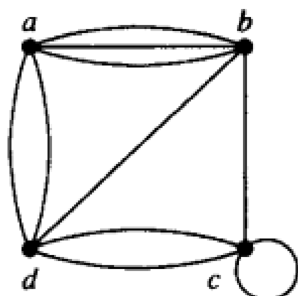


图 8-31 伪图

我们曾在第 7 章里用 0－1 矩阵表示有向图。若有向图 $G=(V, E)$ 从 v_i 到 v_j 有边，则它的矩阵在(i, j)位置上有 1，其中 v_1，v_2，…，v_n 是有向图任意的顶点序列。换句话说，若 $\mathbf{A}=[a_{ij}]$是相对于这个顶点表的邻接矩阵，则

$$a_{ij}=\begin{cases}1, & \text{若}\{v_i,v_j\}\text{ 是 } G \text{ 的一条边} \\ 0, & \text{否则}\end{cases}$$

有向图的邻接矩阵不必是对称的，因为当从 a_i 到 a_j 有边时，从 a_j 到 a_i 可以没有边。

邻接矩阵也可用来表示有向多重图。同样，当有连接两个顶点的同向多重边时，这样的矩阵不是 0－1 矩阵。在有向多重图的邻接矩阵里，a_{ij} 等于关联到(v_i, v_j)的边数。

8.3.4 关联矩阵

表示图的另一种常用方式是用关联矩阵。设 $G=(V, E)$是无向图。设 v_1，v_2，…，v_n 是顶点而 e_1，e_2，…，e_m 是边。则相对于 V 和 E 的这个顺序的关联矩阵是 $n\times m$ 矩阵 $\mathbf{M}=[m_{ij}]$，其中

$$m_{ij}=\begin{cases}1, & \text{当边 } e_j \text{ 关联 } v_i \text{ 时} \\ 0, & \text{否则}\end{cases}$$

例 6 用关联矩阵表示图 8-32 所示的图。

解 关联矩阵是

$$\begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{array} \begin{array}{c} \begin{array}{cccccc} e_1 & e_2 & e_3 & e_4 & e_5 & e_6 \end{array} \\ \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

■

关联矩阵也可用来表示多重图边和环。在关联矩阵里用各项相等的列来表示多重边，因为这些边关联同样一对顶点。用恰有一项等于 1 的列来表示环，它对应于环所关联的顶点。

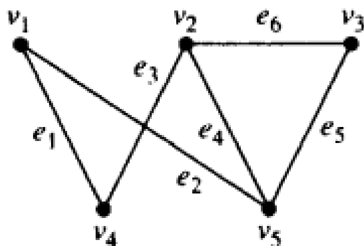


图 8-32 无向图

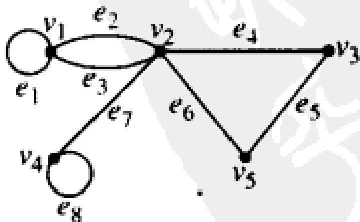


图 8-33 伪图

例 7 用关联矩阵表示图 8-33 所示的伪图。

解 这个图的关联矩阵是

$$\begin{array}{c} \\ v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{array}\begin{array}{c} \begin{array}{cccccccc} e_1 & e_2 & e_3 & e_4 & e_5 & e_6 & e_7 & e_8 \end{array} \\ \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \end{array}$$

■

8.3.5 图的同构

我们经常需要知道是否有可能以同样的方式来画出两个图。例如，在化学里用图为化合物建模。不同的化合物可能分子式相同但结构不同。这样的化合物不能用以同样方式来画的图表示。表示过去已知化合物的图可以用来判定想像中的新化合物是否已经研究过了。

对具有同样结构的图来说，存在着一些有用的术语。

定义 1 设 $G_1=(V_1, E_1)$和 $G_2=(V_2, E_2)$是简单图，若存在一对一的和映上的从 V_1 到 V_2 的函数 f，且 f 具有这样的性质：对 V_1 里所有的 a 和 b 来说，a 和 b 在 G_1 里相邻当且仅当 $f(a)$和 $f(b)$在 G_2 里相邻，就说 G_1 与 G_2 是同构的。这样的函数 f 称为同构⊖。

换句话说，当两个简单图同构时，两个图的顶点之间具有保持相邻关系的一一对应。

例 8 证明图 8-34 所示图 $G=(V, E)$和 $H=(W, F)$同构。

解 函数 f 满足 $f(u_1)=v_1$，$f(u_2)=v_4$，$f(u_3)=v_3$，$f(u_4)=v_2$，它是 V 和 W 之间的一一对应。为了看出这个对应保持相邻关系，注意 G 里相邻的顶点是 u_1 和 u_2，u_1 和 u_3，u_2 和 u_4，以及 u_3 和 u_4，由 $f(u_1)=v_1$ 和 $f(u_2)=v_4$，$f(u_1)=v_1$ 和 $f(u_3)=v_3$，$f(u_2)=v_4$ 和 $f(u_4)=v_2$，以及 $f(u_3)=v_3$ 和 $f(u_4)=v_2$ 所组成的每一对顶点都是在 H 里相邻的。 ■

判定两个简单图是否同构常常是一件困难的事情。在两个带有 n 个顶点的简单图顶点集之间有 $n!$ 种可能的一一对应。若 n 太大，则通过检验每一种对应来看它是否保持相邻关系是不可行的。不过，常常可以通过说明两个简单图不具有同构的图所必须具有的性质来说明它们不同构。这样的性质被称为对简单图的同构来说的不变量。例如，同构的简单图必然具有相同的顶点数，因为在这些图的顶点集之间有着一一对应。而且，同构的简单图必然有相同的边数，因为在顶点之间的一一对应建立了边之间的一一对应。另外，同构的简单图必然有相同的顶点度。即 G 里的 d 度顶点 v 必然对应 H 里的 d 度顶点 $f(v)$，这是因为在 G 里顶点 w 与 v 相邻，当且仅当在 H 里 $f(v)$与 $f(w)$相邻。

例 9 说明图 8-35 所示的图不同构。

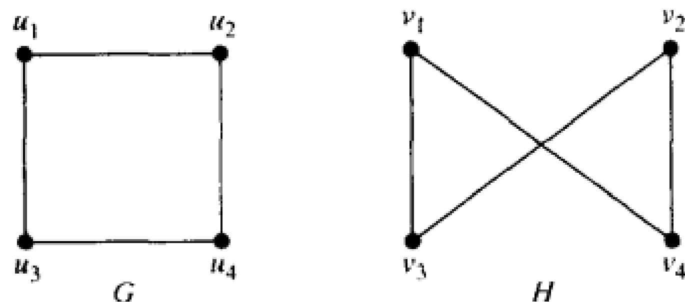


图 8-34 图 G 和 H

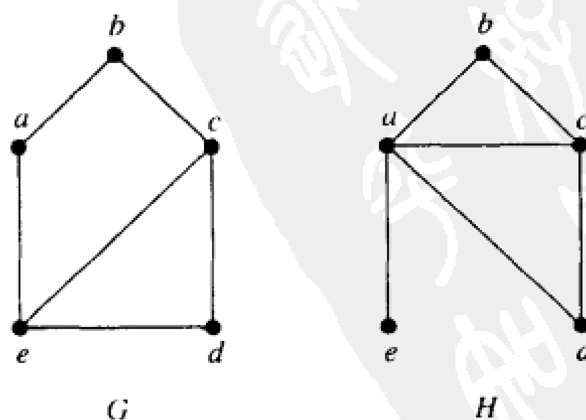


图 8-35 图 G 和 H

⊖ 同构(isomorphism)这个词来自两个希腊语字根：表示“相等”的 isos 和表示“形式”的 morphe。

Extra Examples **解** G 和 H 都具有 5 个顶点 6 条边。不过，G 有 1 度顶点 e 而 H 没有 1 度顶点。所以 G 与 H 是不同构的。■

顶点数、边数以及顶点度都是在同构下的不变量。若两个简单图的这些量有任何不同，则这两个图就不是同构的。不过，当这些量都相同时，也不一定意味着两个图同构。目前还没有已知的用来判定简单图是否同构的不变量集。

例 10 判定图 8-36 所示的图是否同构。

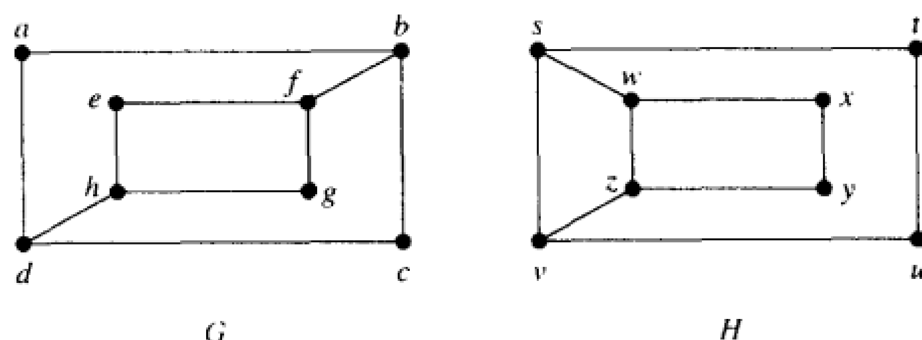


图 8-36 图 G 和 H

解 图 G 和 H 都具有 8 个顶点 10 条边。它们都具有 4 个 2 度顶点和 4 个 3 度顶点。因为这些不变量都相同，所以它们可能会是同构的。

然而 G 和 H 不是同构的。为了看明白这一点，注意到因为在 G 里 $\deg(a)=2$，所以 a 必然对应于 H 里的 t、u、x 或 y，这是因为这些顶点是 H 里的 2 度顶点。不过，H 里的这 4 个顶点中的每一个都与 H 里另一个 2 度顶点相邻，但是在 G 里 a 却不是这样的。

看出 G 与 H 不同构的另一种方式是，注意到若这两个图同构，则由 3 度顶点和连接它们的边所组成的子图同构(读者应当验证它)。然而图 8-37 所示的这些子图却不是同构的。■

为了说明从图 G 的顶点集到图 H 的顶点集的函数 f 是一个同构，需要说明 f 保持边的关系。一种有助于这样做的方式是利用邻接矩阵。具体地说，为了说明 f 是一个同构，就说明 G 的邻接矩阵与 H 的邻接矩阵相同，其中 G 的邻接矩阵的行和列的标记都是 G 里的顶点，H 的邻接矩阵的行和列的标记都是 G 里的对应顶点在 f 下的像。例 11 解释如何这样做。

例 11 判定图 8-38 所示的图 G 和 H 是否同构。

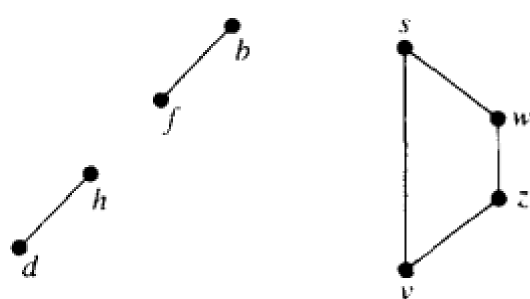


图 8-37 由 3 度顶点和连接它们的边所组成的 G 和 H 的子图

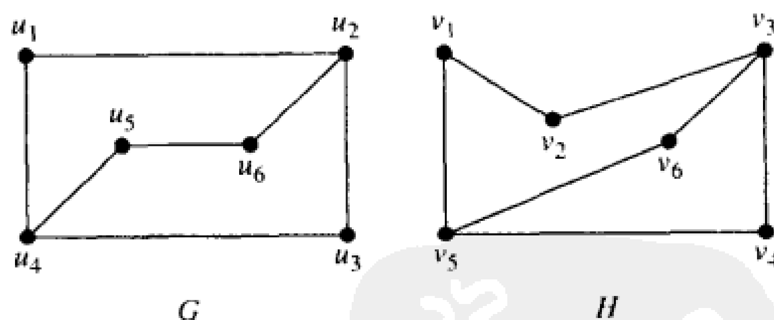


图 8-38 图 G 和 H

解 G 和 H 都具有 6 个顶点 7 条边，都具有 4 个 2 度顶点和 2 个 3 度顶点。还容易看出由 2 度顶点和连接它们的边所组成的 G 和 H 的子图是同构的(读者应当验证它)。因为 G 和 H 对这些不变量来说是相同的，这就有理由试着找出一个同构 f。

现在定义函数 f，然后判定它是否同构。因为 $\deg(u_1)=2$ 而且 u_1 不与任何其他 2 度顶点相邻，所以 u_1 的像必然是 v_4 或 v_6，它们是 H 里仅有的不与 2 度顶点相邻的顶点。任取 $f(u_1)=v_6$。(如果发现这个选择得不出同构，就接着试验 $f(u_1)=v_4$。)因为 u_2 与 u_1 相邻，所以 u_2 可能的像是 v_3

和 v_5。任取 $f(u_2)=v_3$。照这样继续下去，用顶点的相邻关系和度作为指引，令 $f(u_3)=v_4$，$f(u_4)=v_5$，$f(u_5)=v_1$，以及 $f(u_6)=v_2$。现在已经有了在 G 的顶点集与 H 的顶点集之间的一一对应，即 $f(u_1)=v_6$，$f(u_2)=v_3$，$f(u_3)=v_4$，$f(u_4)=v_5$，$f(u_5)=v_1$，以及 $f(u_6)=v_2$。为了查看 f 是否保持边，就检查 G 的邻接矩阵：

$$\mathbf{A}_G = \begin{array}{c} \\ u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \end{array} \begin{array}{c} \begin{array}{cccccc} u_1 & u_2 & u_3 & u_4 & u_5 & u_6 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

以及用 G 里的对应顶点的像来标记行和列的 H 的邻接矩阵：

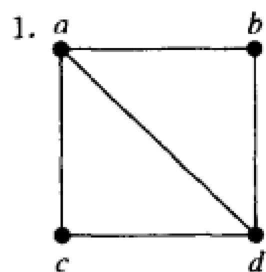
$$\mathbf{A}_H = \begin{array}{c} \\ v_6 \\ v_3 \\ v_4 \\ v_5 \\ v_1 \\ v_2 \end{array} \begin{array}{c} \begin{array}{cccccc} v_6 & v_3 & v_4 & v_5 & v_1 & v_2 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

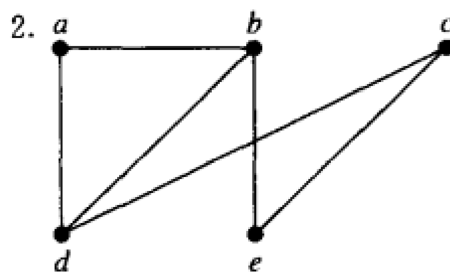
因为 $\mathbf{A}_G=\mathbf{A}_H$，所以 f 是保持边的。由此得出 f 是一个同构，所以 G 与 H 是同构的。注意，若事实证明 f 不是一个同构，是无法得出 G 与 H 不是同构的，因为 G 和 H 里的顶点的另一个对应仍然可以是同构。■

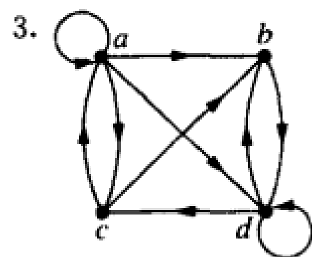
判定两个图是否同构，已知的最好算法具有指数的最坏情形时间复杂性(对图的顶点数来说)。不过，解决这个问题的线性平均情形时间复杂性的算法已经找到，而且有希望找到判定两个图是否同构的多项式最坏情形时间复杂性的算法。一种名为 NAUTY 的最佳实用算法，在现代个人电脑上可在 1 秒钟之内判定带有 100 个顶点的两个图是否同构。可以在因特网上下载 NAUTY 软件并用它做实验。

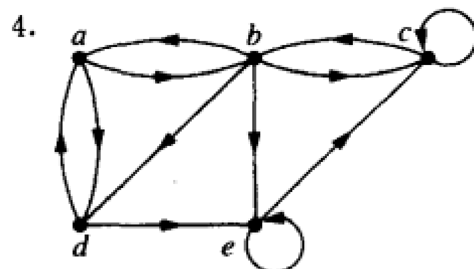
练习

在练习 1～4 中，用邻接表表示给定的图。

1.


2.


3.


4.


5. 用邻接矩阵表示练习 1 里的图。

6. 用邻接矩阵表示练习 2 里的图。

7. 用邻接矩阵表示练习 3 里的图。

8. 用邻接矩阵表示练习 4 里的图。

9. 用邻接矩阵表示下列每一个图。

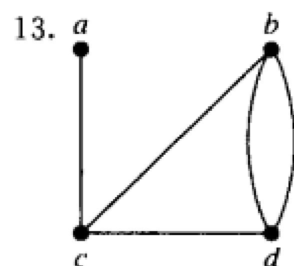
a) K_4　　b) $K_{1,4}$　　c) $K_{2,3}$

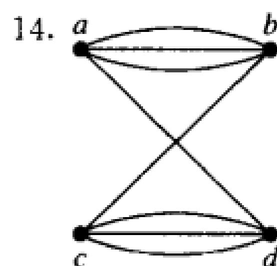
d) C_4　　e) W_4　　f) Q_3

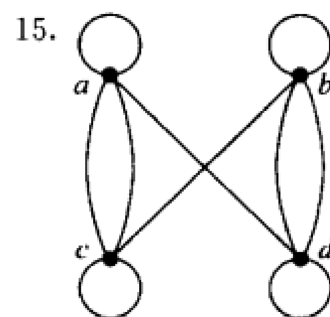
在练习 10～12 中，画出带有给定的邻接矩阵的图。

10. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$　　11. $\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$　　12. $\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}$

在练习 13～15 中，用邻接矩阵表示给定的图。

13. 

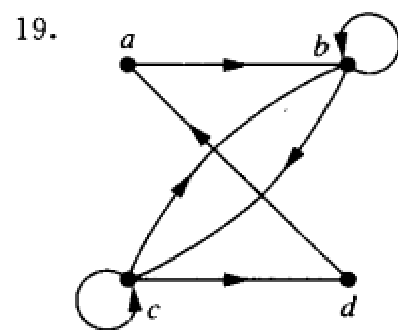
14. 

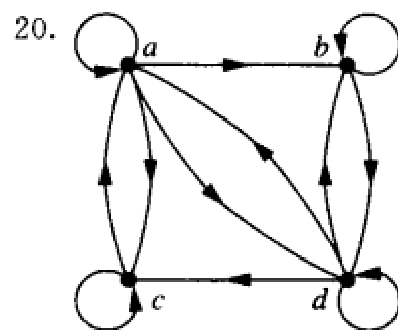
15. 

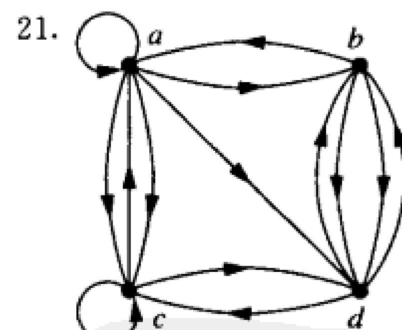
在练习 16～18 里，画出用给定的邻接矩阵所表示的无向图。

16. $\begin{bmatrix} 1 & 3 & 2 \\ 3 & 0 & 4 \\ 2 & 4 & 0 \end{bmatrix}$　　17. $\begin{bmatrix} 1 & 2 & 0 & 1 \\ 2 & 0 & 3 & 0 \\ 0 & 3 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$　　18. $\begin{bmatrix} 0 & 1 & 3 & 0 & 4 \\ 1 & 2 & 1 & 3 & 0 \\ 3 & 1 & 1 & 0 & 1 \\ 0 & 3 & 0 & 0 & 2 \\ 4 & 0 & 1 & 2 & 3 \end{bmatrix}$

在练习 19～21 中，求给定的有向图的邻接矩阵。

19. 

20. 

21. 

在练习 22～24 中，画出用给定的邻接矩阵来表示的图。

22. $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$　　23. $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 0 \\ 0 & 2 & 2 \end{bmatrix}$　　24. $\begin{bmatrix} 0 & 2 & 3 & 0 \\ 1 & 2 & 2 & 1 \\ 2 & 1 & 1 & 0 \\ 1 & 0 & 0 & 2 \end{bmatrix}$

25. 每一个对称的和对角线全为 0 的 0－1 方阵是否都是简单图的邻接矩阵？

26. 用关联矩阵表示练习 1 和练习 2 中的图。

27. 用关联矩阵表示练习 13～15 中的图。

*28. 什么是无向图的邻接矩阵的一行里的各项之和？对有向图来说呢？

*29. 什么是无向图的邻接矩阵的一列里的各项之和？对有向图来说呢？

30. 什么是无向图的关联矩阵的一行里的各项之和？对有向图来说呢？

31. 什么是无向图的关联矩阵的一列里的各项之和？对有向图来说呢？

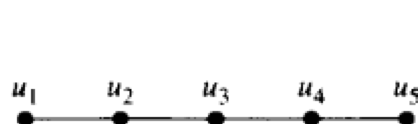
*32. 求下列每个图的邻接矩阵。

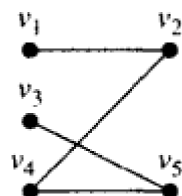
a) K_n b) C_n c) W_n d) $K_{m,n}$ e) Q_n

*33. 求练习 32(a)～(d)中的图的关联矩阵。

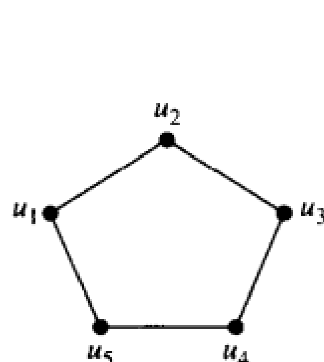
在练习 34～44 中，判定所给定的一对图是否同构。给出一个同构或给出不存在同构的严格证明。

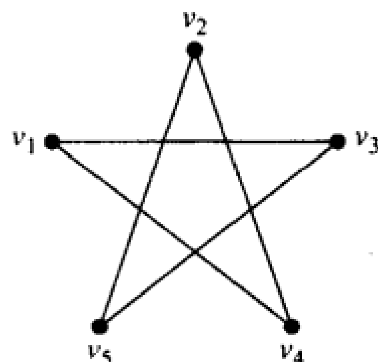
34.



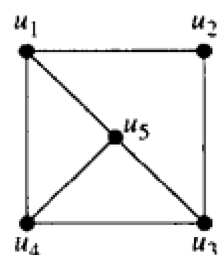


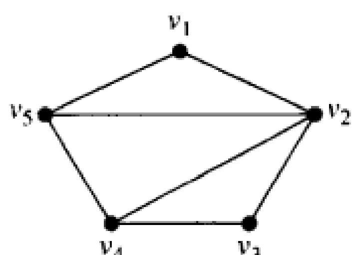
35.



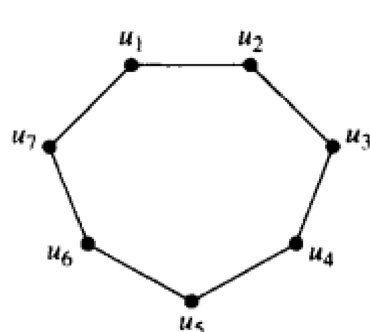


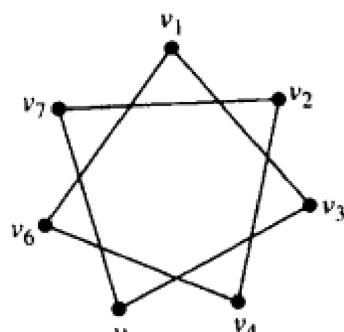
36.



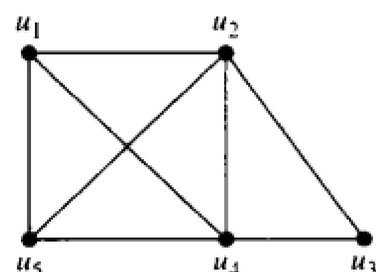


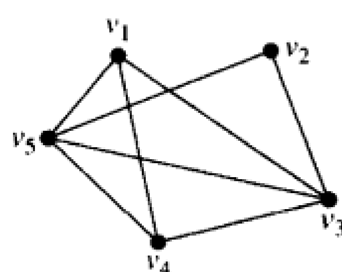
37.



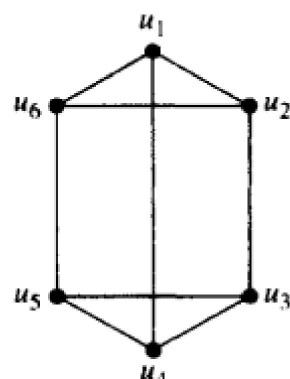


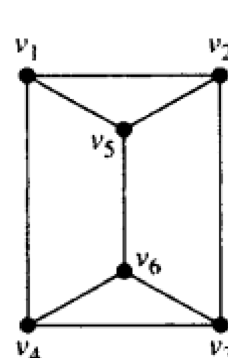
38.



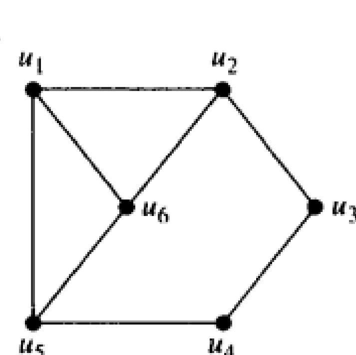


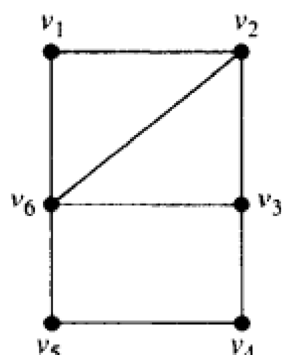
39.



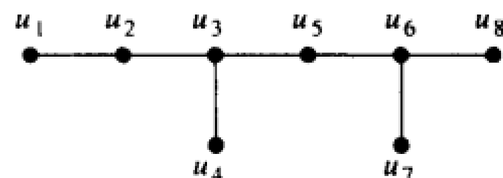


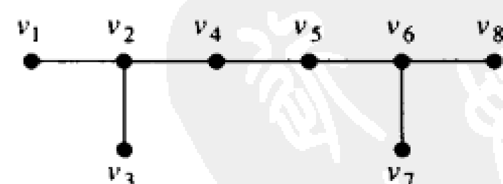
40.



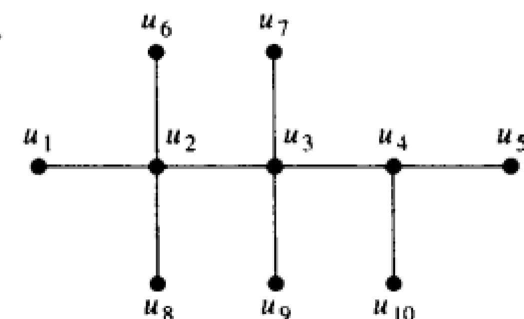


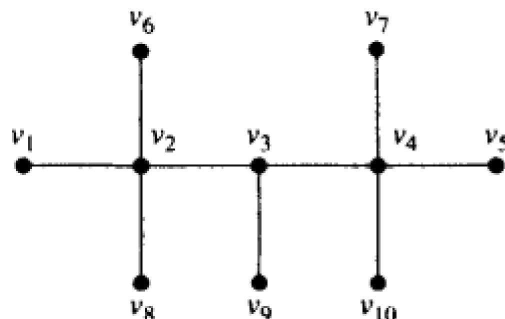
41.



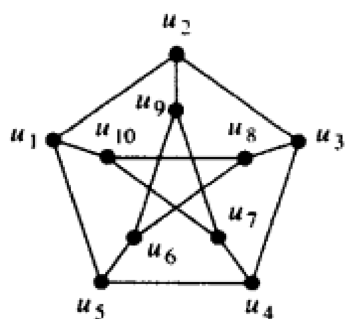


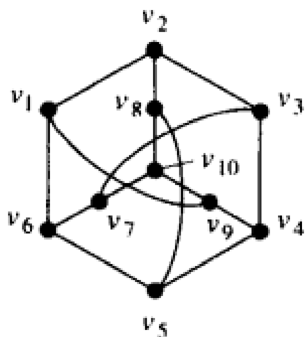
42.



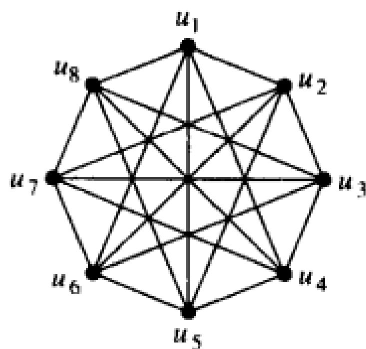


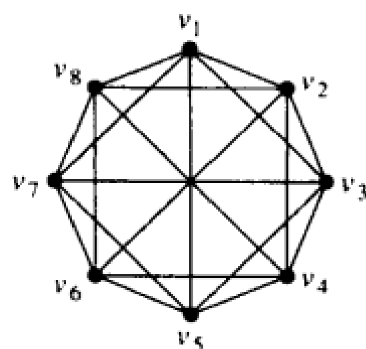
43.





44.





45. 证明：简单图的同构关系是等价关系。

46. 设 G 和 H 是同构的简单图。证明：它们的补图 $\overline{G}$ 和 $\overline{H}$ 也是同构的。

47. 描述一下图的邻接矩阵里对应于孤立点的行和列。

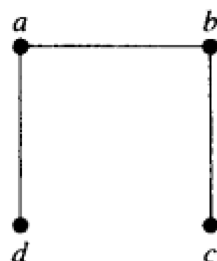
48. 描述一下图的关联矩阵里对应于孤立点的行。

49. 证明：带有 2 个以上顶点的偶图的顶点可以排序，使得邻接矩阵形如

$$\begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{B} & \mathbf{0} \end{bmatrix}$$

其中所示的四项都是矩形块。

若简单图 G 和 $\overline{G}$ 是同构的，则 G 称为自补图。



50. 证明：右图是自补图。

51. 求带有 5 个顶点的自补简单图。

*52. 证明：若 G 是带有 v 个顶点的自补简单图，则 $v \equiv 0$ 或 $1 \pmod 4$。

53. 对哪些整数 n 来说 C_n 是自补图？

54. 带有 n 个顶点的非同构的简单图有多少个？其中 n 是

a)2 b)3 c)4

55. 带有 5 个顶点和 3 条边的非同构的简单图有多少个？

56. 带有 6 个顶点和 4 条边的非同构的简单图有多少个？

57. 带有下列邻接矩阵的简单图是否同构？

a) $\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$，$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

b) $\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$，$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$

c) $\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$，$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$

58. 判定带有下列关联矩阵的无环图是否同构。

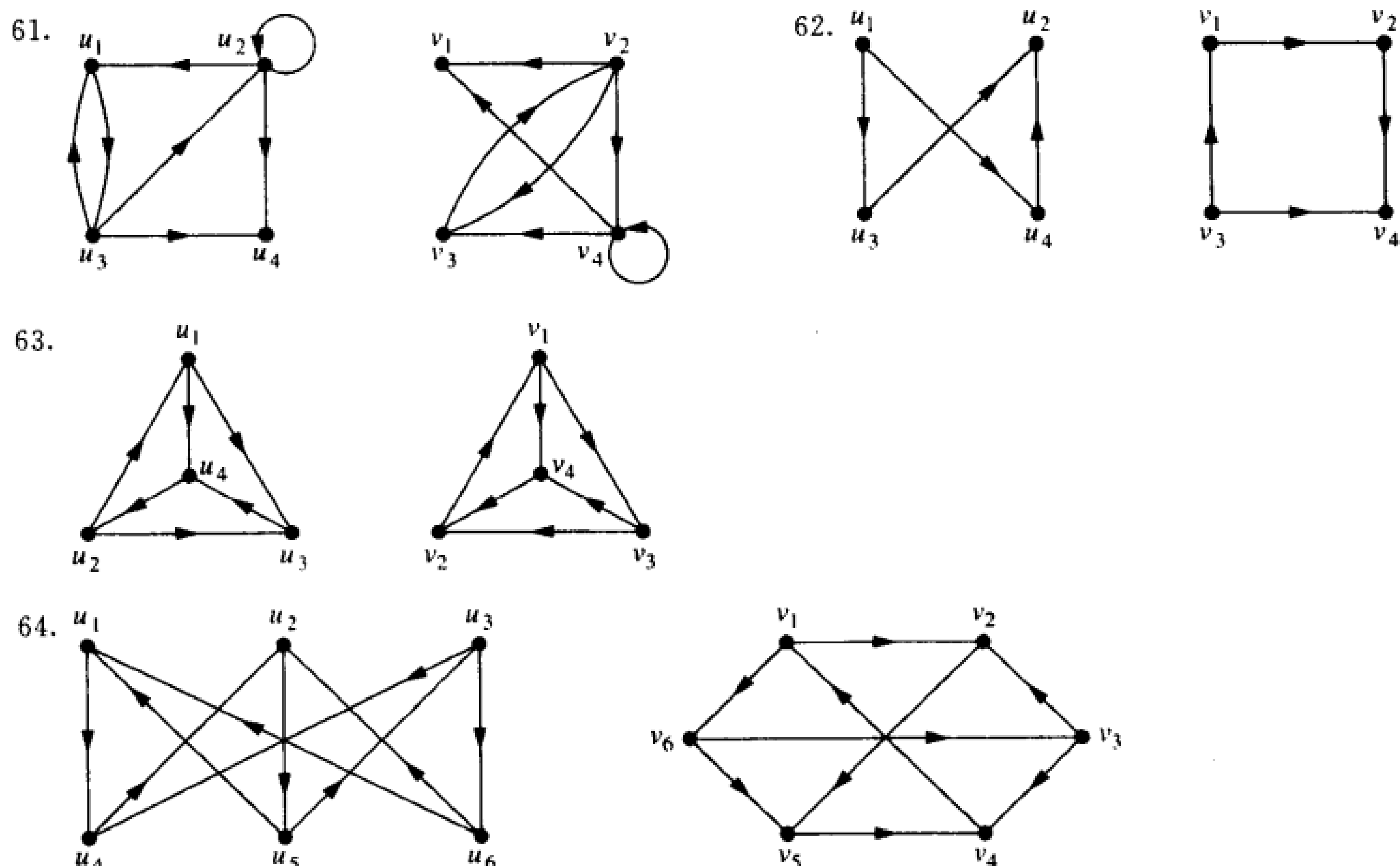
a) $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$，$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$

b) $\begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{bmatrix}$，$\begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}$

59. 把简单图的同构定义推广到包含环和多重边的无向图。

60. 定义有向图的同构。

在练习 61～64 里，判定所给定的一对图是否同构。



65. 证明：若 G 和 H 是同构的有向图，则 G 和 H 的逆(在8.2节练习47之前定义)也是同构的。

*66. 带有 n 个顶点的非同构有向图有多少个？其中 n 是

a)2　　b)3　　c)4

*67. 无向图的关联矩阵与它的转置之积是什么？

*68. 表示带有 v 个顶点和 e 条边的简单图需要多少存储空间？其中分别利用

a)邻接表　　b)邻接矩阵　　c)关联矩阵

魔鬼对是一对不同构的图，但所谓的同构检验不能证明其不同构。

69. 求一个用于检验的魔鬼对，检查两个图里的顶点度的序列，确定它们是相同的。

8.4 连通性

8.4.1 引言

许多问题可以用沿图的边旅行所形成的通路来建模。例如，判定在两个计算机之间用中间连接能否传递消息的问题，就可以用图模型来研究。利用图里的通路所组成的模型，可以解决投递邮件、收取垃圾以及计算机网络诊断等有效规划路线的问题。

8.4.2 通路

非形式化地说，通路是边的序列，它从图的一个顶点上开始并且行经图的边，总是连接相邻的顶点对。定义1给出通路的形式化定义和相关术语。

定义1 设 n 是非负整数且 G 是无向图。在 G 中从 u 到 v 的长度为 n 的通路是 G 的 n 条边 e_1，…，e_n 的序列，使得 $f(e_1)=\{x_0, x_1\}$，$f(e_2)=\{x_1, x_2\}$，…，$f(e_n)=\{x_{n-1}, x_n\}$，其中 $x_0=u$ 而 $x_n=v$。当这个图是简单图时，就用顶点序列 x_0，x_1，…，x_n 表示这条通路(这是因为列出这些顶点就唯一地确定了通路)。若一条通路在相同的顶点上开始和结束，即 $u=v$ 且长度大于0，则它是一条回路。把这条通路或回路说成是经过顶点 x_1，x_2，…，x_{n-1} 或遍历边 e_1，e_2，…，

e_n。若通路或回路不重复地包含相同的边，则它是简单的。

当没有必要区分多重边时，就用顶点序列 x_0，x_1，…，x_n 表示通路 e_1，e_2，…，e_n，其中对 $i=1$，2，…，n 来说，$f(e_i)=\{x_{i-1}, x_i\}$。这种记法仅仅指出通路所经过的顶点。经过这个顶点序列可以有多条通路。注意长度为 0 的通路由单个顶点组成。

例 1 在图 8-39 所示的图里，a，d，c，f，e 是长度为 4 的简单通路，因为$\{a, d\}$，$\{d, c\}$，$\{c, f\}$和$\{f, e\}$都是边。但是 d，e，c，a 不是通路，因为$\{e, c\}$不是边。注意 b，c，f，e，b 是长度为 4 的回路，这是因为$\{b, c\}$，$\{c, f\}$，$\{f, e\}$和$\{e, b\}$都是边，这条通路在 b 上开始和结束。长度为 5 的通路 a，b，e，d，a，b 不是简单的，因为它两次包含边$\{a, b\}$。■

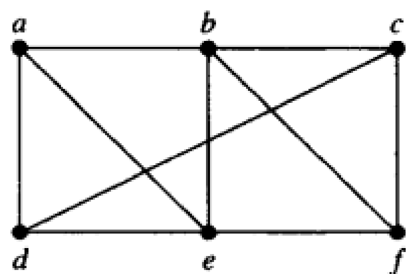


图 8-39 简单图

有向图里的通路和回路在第 7 章里介绍过。现在对有向多重图来定义这样的通路。

定义 2 设 n 是非负整数且 G 是有向多重图。在 G 中从 u 到 v 的长度为 n 的**通路**是 G 的边的序列 e_1，e_2，…，e_n，使得 $f(e_1)=(x_0, x_1)$，$f(e_2)=(x_1, x_2)$，…，$f(e_n)=(x_{n-1}, x_n)$，其中 $x_0=u$ 而 $x_n=v$。当有向图里没有多重边时，就用顶点序列 x_0，x_1，…，x_n 表示这条通路。把在相同的顶点上开始和结束的长度大于 0 的通路称为**回路**或**圈**。若一条通路或回路不重复地包含相同的边，则把它称为**简单的**。

注意通路上一条边的终点是这条通路上下一条边的起点。当没有必要区分多重边时，就用顶点序列 x_0，x_1，…，x_n 表示通路 e_1，e_2，…，e_n，其中对 $i=1$，2，…，n 来说，$f(e_i)=(x_{i-1}, x_i)$。这种记法仅仅指出通路所经过的顶点。可以有多条通路经过这个顶点序列。

通路表示许多图模型里的有用信息，如例题 2～4 所示。

例 2 熟人关系图中的通路 在熟人关系图中，如果相邻两个人彼此认识则将此二人连接起来，形成一条人链。例如在 8.1 节的图 8-7 中，有一条连接 Kamini 和 Ching 的 6 个人的链。许多社会学家猜想，是否可以用只包含 5 个或更少的人的短链来连接世界上几乎每一对人。这意味着世界上所有人的熟人关系图中几乎每对顶点都通过长度不超过 4 的通路来连接。约翰·奎尔(John Guare)的戏剧《六度分离(Six Degrees of Separation)》就是基于这个概念。■

例 3 合作图中的通路 在合作图中，如果表示作者的两个顶点 a 和 b 之间有从 a 开始到 b 结束的一系列作者，使得每条边端点所表示的两个作者写过一篇联名论文，则 a 和 b 就通过一条通路而连接。在所有数学家的合作图中，数学家 m 的埃德斯数(在第 7 章补充练习 14 中用关系术语定义过)就是在 m 和代表成果极其丰富的数学家保罗·埃德斯(1996 年去世)的顶点之间的最短通路的长度。换句话说，一个数学家的埃德斯数就是从保罗·埃德斯开始到这个数学家结束的最短的数学家链的长度，其中每一对相邻的数学家都联名写过论文。根据“埃德斯数项目”，2001 年 11 月份具有各个埃德斯数的数学家的数目如表 8-4 所示。■

例 4 好莱坞图中的通路 在好莱坞图(参见 8.1 节例 4)中，当存在连接两个顶点 a 和 b 的演员链，其中在这个链上每两个相邻的演员都出演过同一部电影时，a 和 b 就被连接。在好莱坞图中演员 c 的培根数定义为连接 c 和著名演员凯文·培根的最短通路的长度。随着新电影(包括凯文·培根的新电影)的不断产生，演员们的培根数也在不断地发生变化。表 8-5 中显示的是从培根网站得来的 2001 年 11 月份具有各个培根数的演员的数目。■

表 8-4 具有给定埃德斯数的数学家的数目

埃德斯数	人数	埃德斯数	人数
0	1	8	3 214
1	502	9	953
2	713	10	262
3	26 422	11	94
4	62 136	12	23
5	66 157	13	4
6	32 280	14	7
7	10 431	15	1

表 8-5 具有给定培根数的演员数

培根数	人数	培根数	人数
0	1	6	534
1	1 479	7	81
2	115 204	8	28
3	285 929	9	1
4	65 021	10	1
5	4 535		

8.4.3 无向图的连通性

若消息通过一个或多个中间计算机来传递，则计算机网络何时具有每对计算机都可共享信息的性质？当利用图来表示这个计算机网络，用顶点表示计算机而用边表示通信连接时，这个问题就变成：何时在图里任何两个顶点之间都存在通路？

定义 3 若无向图每一对不同的顶点之间都有通路，则该图称为**连通的**。

因此，在网络里任何两个计算机都可以通信，当且仅当这个网络图是连通的。

例 5 图 8-40 里的图 G_1 是连通的，这是因为在每一对不同的顶点之间都有通路(读者应当验证它)。但是图 8-40 里的图 G_2 不是连通的。例如，在顶点 a 和 d 之间没有通路。■

在第 9 章里将用到下述定理。

定理 1 在连通无向图的每一对不同顶点之间都存在简单通路。

证 设 u 和 v 是连通无向图 $G=(V, E)$ 的两个不同的顶点。因为 G 是连通的，所以 u 和 v 之间至少有 1 条通路。设 x_0，x_1，…，x_n 是长度最短的通路的顶点序列，其中 $x_0=u$ 而 $x_n=v$。这条长度最短的通路是简单的。为了看明白这一点，假设它不是简单的。则对满足 $0 \leqslant i<j$ 的某个 i 和 j 来说，有 $x_i=x_j$。这意味着通过删除顶点序列 x_i，…，x_{j-1} 所对应的边，就获得了带有顶点序列 x_0，x_1，…，x_{i-1}，x_j，…，x_n 的从 u 到 v 的更短的通路。□

不连通的图是 2 个或 2 个以上连通子图之并，每一对子图都没有公共的顶点。这些不相交的连通子图称为图的**连通分支**。

例 6 图 8-41 所示的图 H 的连通分支是什么？

解 图 H 是图 8-41 所示的 3 个不相交连通子图 H_1、H_2 和 H_3 之并。这 3 个子图是 H 的连通分支。■

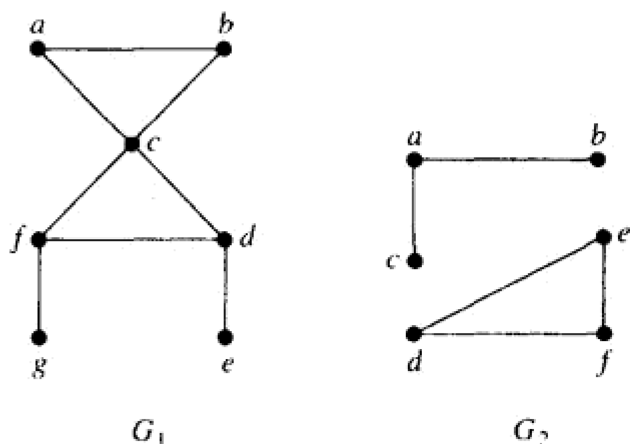


图 8-40 图 G_1 和 G_2

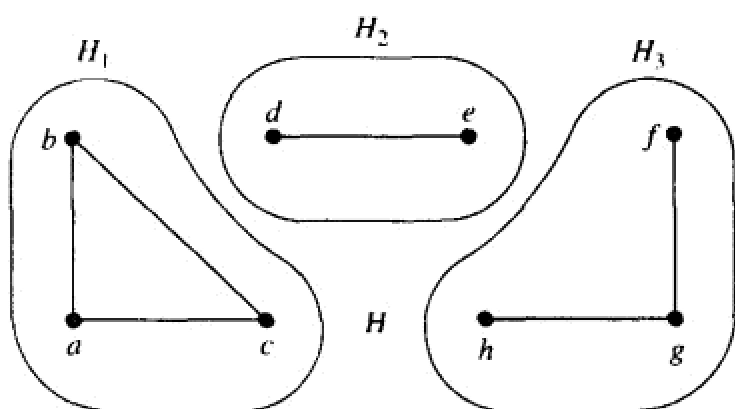


图 8-41 图 H 和它的连通分支 H_1，H_2 和 H_3

例 7 呼叫图的连通分支 当电话呼叫图(参见 8.1 节例 7)中存在一系列从 x 开始到 y 结束的电话呼叫时，两个顶点 x 和 y 就属于同一个连通分支。当分析 AT&T 网络中特定一天内发生的电话呼叫的呼叫图时，发现这个图具有 53 767 087 个顶点、超过 1 亿 7 千万条边和超过 370 万个连通分支。这些连通分支大多数都很小：大约四分之三是由表示只在彼此之间呼叫的一对电话号码的两个顶点所组成。这个图具有一个包含 44 989 297 个顶点(占总数的 80%)的巨大的连通分支。另外，这个连通分支中每个顶点都通过一条不超过 20 个顶点的链连接到任何其他顶点。 ■

有时删除一个顶点和它所关联的边，就产生带有比原图更多的连通分支的子图。把这样的顶点称为割点(或节点)。从连通图里删除割点，就产生不连通的子图。同理，把一旦删除就产生带有比原图更多的连通分支的子图的边称为割边或桥。

例 8 求出图 8-42 所示的图 G 的割点和割边。

解 图 G 的割点是 b、c 和 e。删除这些顶点中的一个(和它的邻边)，就使得这个图不再是连通的。割边是 $\{a, b\}$ 和 $\{c, e\}$。删除这些边中的一条，就使得 G 不再是连通的。 ■

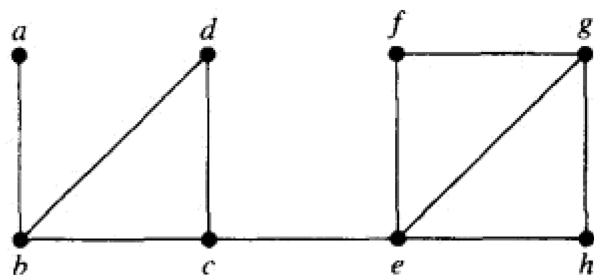


图 8-42 图 G

8.4.4 有向图的连通性

根据是否考虑边的方向，在有向图里有两种连通性概念。

定义 4 若每当 a 和 b 都是一个有向图里的顶点时，就有从 a 到 b 和从 b 到 a 的通路，则该图是强连通的。

为了让一个有向图是强连通的，从这个图里任何一个顶点到任何另外一个顶点必须存在有向边的序列。有向图可以不是强连通的但还是"一整块"。定义 5 准确地说明了这个概念。

定义 5 若在有向图的底图里，任何两个顶点之间都有通路，则该有向图是弱连通的。

有向图是弱连通的，当且仅当在忽略边的方向时，任何两个顶点之间总是存在通路。显然，任何强连通有向图也是弱连通的。

例 9 图 8-43 所示的有向图 G 和 H 是否为强连通的？是否为弱连通的？

解 G 是强连通的，因为在这个有向图里，任何两个顶点之间都存在通路(读者应当验证它)。

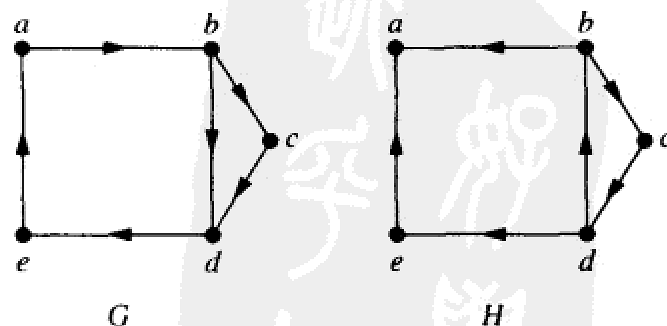


图 8-43 有向图 G 和 H

因此 G 也是弱连通的。图 H 不是强连通的，在这个图里，从 a 到 b 没有有向通路。但是 H 是弱连通的，因为在 H 的底图里，任何两个顶点之间都有通路(读者应当验证它)。 ■

有向图 G 的子图是强连通的而不包含在更大的强连通子图中，即极大强连通子图，可称之为 G 的**强连通分支**或**强分支**。

例 10 图 8-43 中的图 H 有 3 个强连通分支，包括：顶点 a；顶点 e；由顶点 b、c 和 d 以及边(b, c)、(c, d)和(d, b)所组成的图。 ■

例 11 网络图的强连通分支 8.1 节例 8 介绍的网络图用顶点表示网页并且用有向边表示链接。网络在 1999 年的快照产生了具有 2 亿个顶点和 15 亿条边的网络图(参见[Br00]来了解详情)。网络图的底无向图不是连通的并且有一个连通分支包含了这个图中大约 90%的顶点。底无向图的这个连通分支所对应的原来有向图的子图(即具有相同的顶点以及连接这个图中顶点的所有有向边)有一个非常大的强连通分支和许多小的强连通分支。前者称为这个有向图的**巨型强连通分支**。可以遵循从这个分支里任何其他网页开始的链接来到达这个分支里的某一个网页。已经发现这项研究产生的网络图里的巨型强连通分支有超过 5300 万个顶点。这个无向图的大型连通分支里的其余顶点表示 3 种不同类型的网页：可以从巨型强连通分支里的网页来到达，但是不能通过一系列链接来返回前面这些网页的网页；可以通过一系列链接来返回巨型强连通分支里的网页，但是不能通过巨型强连通分支里网页上的链接来到达的网页；既不能到达巨型强连通分支里的网页，也不能通过一系列链接从巨型强连通分支里网页来到达的网页。在这项研究里，发现其余这三个集合中每个都具有大约 4400 万个顶点(这三个集合都接近同样的规模，这是相当令人惊讶的)。 ■

8.4.5 通路与同构

有多种方式可以利用通路和回路来帮助判定两个图是否同构。例如，具有特定长度的简单回路的存在性，就是一种可以用来证明两个图是不同构的有用的不变量。另外，可以利用通路来构造可能的同构映射。

前面指出过，简单图的一个有用的同构不变量是长度为 k 的简单回路的存在性，其中 k 是大于 2 的正整数(这个不变量的证明在本节末尾留作练习 44)。例 12 说明如何用这个不变量来证明两个图不同构。

例 12 判定图 8-44 所示的图 G 和 H 是否同构。

解 G 和 H 都具有 6 个顶点和 8 条边。各自具有 4 个 3 度顶点和 2 个 2 度顶点。所以对两个图来说有 3 个不变量(顶点数、边数以及顶点度)都是相同的。但是 H 有长度为 3 的简单回路，即 v_1，v_2，v_6，v_1，而 G 没有长度为 3 的简单回路(G 里的所有简单回路的长度至少为 4)。因为长度为 3 的简单回路的存在性是一个同构不变量，所以 G 和 H 不同构。 ■

我们已经说明了如何用某种类型的通路，即具有特定长度的简单回路，来证明两个图是不同构的。还可以用通路来求出潜在的同构映射。

例 13 判定图 8-45 所示的图 G 和 H 是否同构。

解 G 和 H 都具有 5 个顶点和 6 条边，都具有 2 个 3 度顶点和 3 个 2 度顶点，都具有 1 个长度为 3 的简单回路，1 个长度为 4 的简单回路，以及 1 个长度为 5 的简单回路。因为所有这些同构不变量都是相同的，所以 G 和 H 可能是同构的。为了求出可能的同构，沿着经过所有顶点并且使得两图里的对应顶点的度都相同的通路前进。例如，G 里的通路 u_1，u_4，u_3，u_2，u_5 和 H 里的通路 v_3，v_2，v_1，v_5，v_4 都经过图里的每一个顶点，都在 3 度顶点上开始，都分别经过 2 度顶点、3 度顶点和 2 度顶点并且在 2 度顶点上结束。通过在图里沿着这些通路前进，定义映射 f

满足$f(u_1)=v_3$，$f(u_4)=v_2$，$f(u_3)=v_1$，$f(u_2)=v_5$ 和 $f(u_5)=v_4$。通过说明 f 保持边，或者通过说明在顶点的适当顺序下 G 和 H 的邻接矩阵是相同的，读者就可以说明 f 是一个同构，所以 G 与 H 是同构的。■

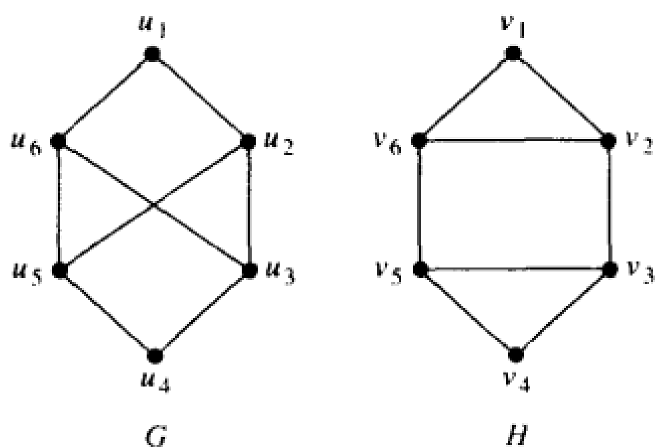


图 8-44 图 G 和 H

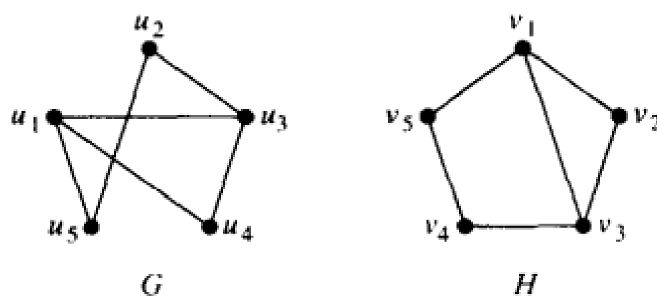


图 8-45 图 G 和 H

8.4.6 计算顶点之间的通路数

在一个图里两个顶点之间的通路的数目，可以用这个图的邻接矩阵来确定。

定理 2 设 G 是带有相对于顶点顺序 v_1，v_2，…，v_n 的邻接矩阵 $\mathbf{A}$ 的图(允许带有无向或有向边、带有多重边和环)。从 v_i 到 v_j 的长度为 r 的不同通路的数目等于 $\mathbf{A}^r$ 的第(i, j)项，其中 r 是正整数。

证 用数学归纳法证明。设 G 是带有邻接矩阵 $\mathbf{A}$ 的图(假设 G 的顶点具有顺序 v_1，v_2，…，v_n)。从 v_i 到 v_j 长度为 1 的通路数是 $\mathbf{A}$ 的第(i, j)项，这是因为该项是从 v_i 到 v_j 的边数。

假设 $\mathbf{A}^r$ 的第(i, j)项是从 v_i 到 v_j 的长度为 r 的不同通路的个数。这是归纳假设。因为 $\mathbf{A}^{r+1}=\mathbf{A}^r\mathbf{A}$，所以 $\mathbf{A}^{r+1}$ 的第(i, j)项等于

$$b_{i1}a_{1j}+b_{i2}a_{2j}+\cdots+b_{in}a_{nj}$$

其中 b_{ik} 是 $\mathbf{A}^r$ 的第(i, k)项。根据归纳假设，b_{ik} 是从 v_i 到 v_k 的长度为 r 的通路数。

从 v_i 到 v_j 的长度为 $r+1$ 的通路，包括从 v_i 到某个中间顶点 v_k 的长度为 r 的通路以及从 v_k 到 v_j 的边。根据计数的乘法规则，这样的通路的个数是从 v_i 到 v_k 的长度为 r 的通路数(即 b_{ik})与从 v_k 到 v_j 的边数(即 a_{kj})之积。当对所有可能的中间顶点 v_k 求这些乘积之和时，根据计数的加法规则，就得出所需要的结果。□

例 14 在图 8-46 所示的简单图 G 里，从 a 到 d 的长度为 4 的通路有多少条？

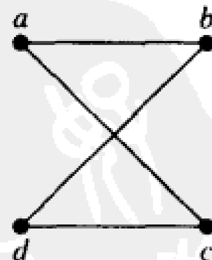


图 8-46 图 G

解 G 的邻接矩阵(顶点顺序为 a，b，c，d)是

$$\mathbf{A}=\begin{bmatrix}0&1&1&0\\1&0&0&1\\1&0&0&1\\0&1&1&0\end{bmatrix}$$

因此从 a 到 d 的长度为 4 的通路数是 $\mathbf{A}^4$ 的第(1, 4)项。因为

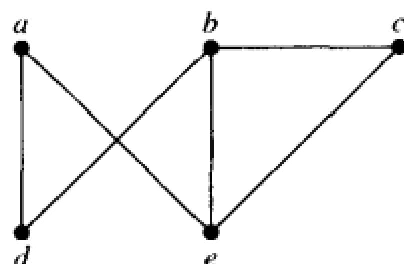
$$\mathbf{A}^4=\begin{bmatrix}8&0&0&8\\0&8&8&0\\0&8&8&0\\8&0&0&8\end{bmatrix}$$

Extra Examples 所以恰好有 8 条从 a 到 d 的、长度为 4 的通路。通过检查这个图，我们看出 a，b，a，b，d；a，b，a，c，d；a，b，d，b，d；a，b，d，c，d；a，c，a，b，d；a，c，a，c，d；a，c，d，b，d 和 a，c，d，c，d 是 8 条从 a 到 d 的通路。■

定理 2 可以用来求出在图的两个顶点之间的最短通路的长度(见练习 40)，还可以用来判定图是否连通(见练习 45 和 46)。

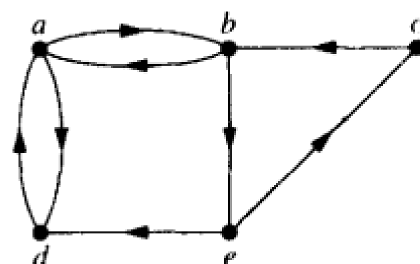
练习

1. 下述的每个顶点列表是否可以构成右图里的通路？哪些通路是简单的？哪些是回路？那些通路的长度是多少？



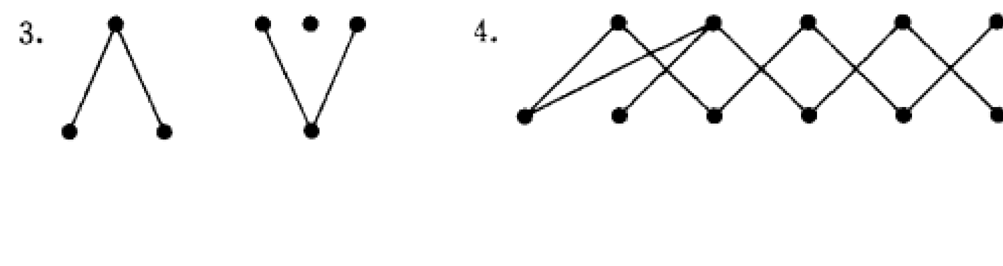
   a) a，e，b，c，b
   b) a，e，a，d，b，c，a
   c) e，b，a，d，b，e
   d) c，b，d，a，e，c

2. 下述的每个顶点列表是否可以构成右图里的通路？哪些通路是简单的？哪些是回路？那些通路的长度是多少？



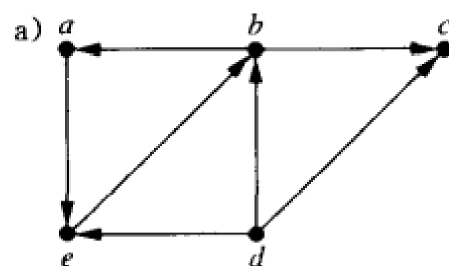
   a) a，b，e，c，b
   b) a，d，a，d，a
   c) a，d，b，e，a
   d) a，b，e，c，b，d，a

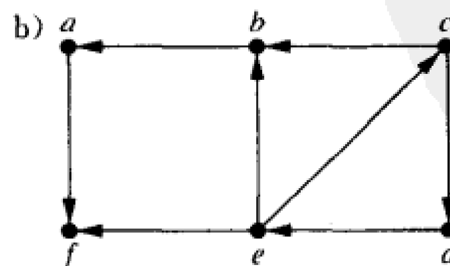
在练习 3～5 里，判定所给的图是否连通。



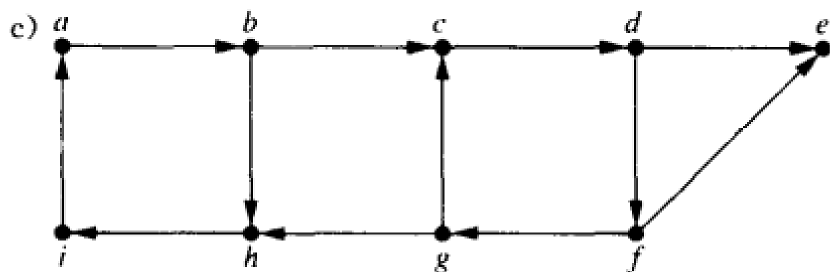
5.

6. 在练习 3～5 里，每个图各自有多少个连通分支？对每个图求出它的每个连通分支。
7. 熟人关系图的连通分支表示什么？
8. 合作图的连通分支表示什么？
9. 解释一下为什么在数学家的合作图里，表示一个数学家的顶点与表示保罗·埃德斯的顶点是在同一个连通分支里，当且仅当这个数学家具有有穷的埃德斯数。
10. 在好莱坞图里(参见 8.1 节例 4)，什么时候表示一个演员的顶点与表示凯文·培根的顶点是在同一个连通分支里？
11. 电话呼叫图的强连通分支表示什么？
12. 求下列每个图的强连通分支。



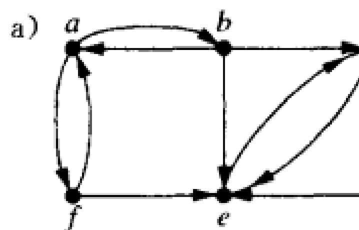


c)

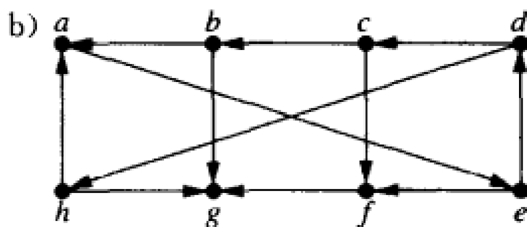


13. 求下列每个图的强连通分支。

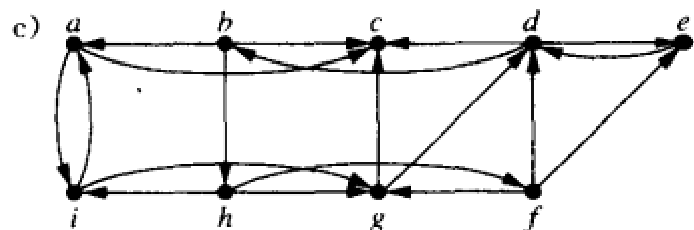
a)



b)



c)



14. 证明：连接有向图同一个强连通分支里两个顶点的有向通路所访问的所有顶点也都在这个强连通分支里。
15. 求出 K_4 里两个不同顶点之间长度为 n 的通路的数目，若 n 是

a)2　　b)3　　c)4　　d)5

16. 对练习 15 里的 n 值来说，求出 $K_{3,3}$ 里任意两个相邻顶点之间长度为 n 的通路的数目。
17. 对练习 15 里的 n 值来说，求出 $K_{3,3}$ 里任意两个不相邻顶点之间长度为 n 的通路的数目。
18. 求出在图 8-39 中 c 和 d 之间具有如下长度的通路的数目：

a)2　　b)3　　c)4

d)5　　e)6　　f)7

19. 求出在练习 2 里的有向图中从 a 到 e 具有如下长度的通路的数目：

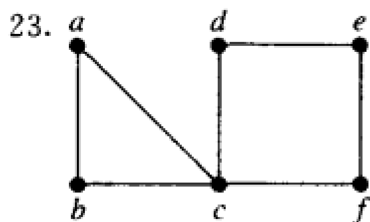
a)2　　b)3　　c)4

d)5　　e)6　　f)7

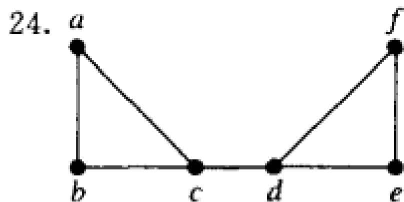
*20. 证明：带有 n 个顶点的连通图至少具有 $n-1$ 条边。
21. 设 $G=(V, E)$ 是简单图。设 R 是 V 上的关系，它是由这样的顶点对 (u, v) 所组成的，使得存在从 u 到 v 的通路或使得 $u=v$。证明：R 是一个等价关系。
*22. 证明：在任何简单图里，任何奇数度顶点之间都有通路。

在练习 23～25 里，求所给图的所有割点。

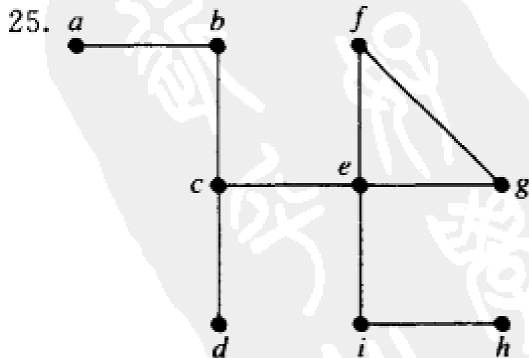
23.



24.



25.



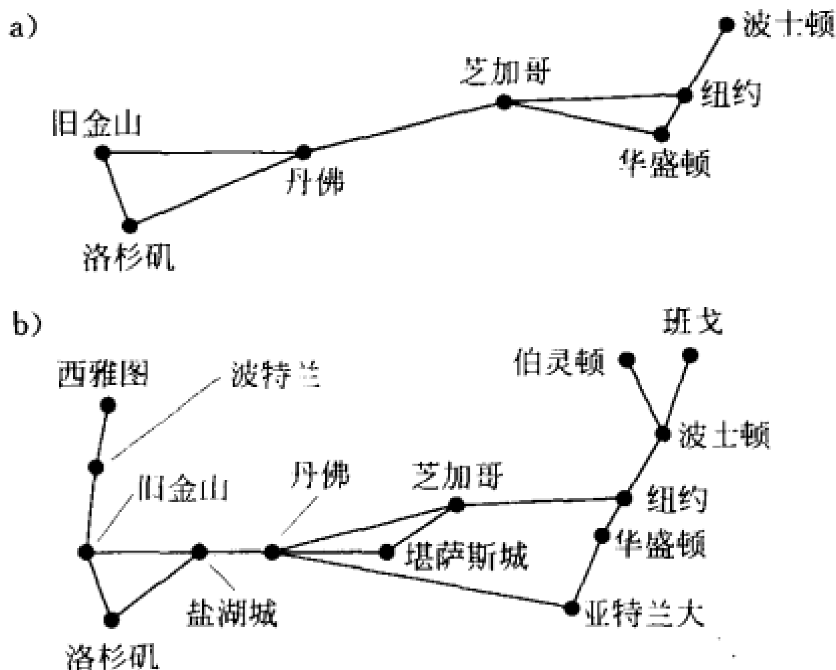
26. 求练习 23～25 里的图的所有割边。
*27. 假设 v 是一条割边的端点。证明：v 是割点当且仅当它不是悬挂点。
*28. 证明：在连通简单图 G 里，顶点 c 是割点当且仅当存在着与 c 不同的顶点 u 和 v，使得在 u 和 v 之间的

每一条通路都经过 c。

*29. 证明：在至少有 2 个顶点的简单图中，至少有 2 个顶点不是割点。

*30. 证明：简单图里的一条边是割边，当且仅当它不属于该图里任何一条简单回路。

31. 若网络里的通信连接故障会导致不能传送某些消息，则应当提供备份连接。对下面(a)和(b)里所示的通信网络来说，确定出那些应当有备份的连接。



一个有向图里的**顶点基**是顶点的一个集合，使得到该有向图里该集合之外的任何顶点，都有从该集合里某个顶点开始的一条通路，并且该集合里任何两个顶点之间都没有通路。

32. 对 8.2 节练习 7～9 里的每个有向图来说，求顶点基。

33. 在影响图(在 8.1 节例 3 里描述)里顶点基的重要性是什么？找出在该例里的影响图的顶点基。

☞34. 证明：若连通有向图 G 是图 G_1 和 G_2 的并图，则 G_1 和 G_2 至少具有 1 个公共的顶点。

*35. 证明：若简单图 G 有 k 个连通分支，而且这些分支分别具有 n_1，n_2，…，n_k 个顶点，则 G 的边数不超过

$$\sum_{i=1}^{k} C(n_i, 2)$$

*36. 用练习 35 证明：带有 n 个顶点和 k 个连通分支的简单图最多有 $(n-k)(n-k+1)/2$ 条边。[提示：先证明

$$\sum_{i=1}^{k} n_i^2 \leqslant n^2-(k-1)(2n-k)$$

其中 n_i 是第 i 个连通分支的顶点数。]

*37. 证明：若带有 n 个顶点的简单图 G 具有超过 $(n-1)(n-2)/2$ 条边，则它是连通的。

38. 当把每个连通分支里的顶点都连续地列出时，描述一下带有 n 个连通分支的图的邻接矩阵。

39. 当 n 取如下值时，存在多少个不同构的、带有 n 个顶点的连通简单图？

a)2　　b)3　　c)4　　d)5

40. 解释一下如何用定理 2 求出图里从顶点 v 到顶点 w 的最短通路的长度。

41. 用定理 2 求出图 8-39 里的多重图从 a 到 f 的最短通路的长度。

42. 用定理 2 求出练习 2 里的有向图从 a 到 c 的最短通路的长度。

☞43. 设 P_1 和 P_2 是简单图 G 里顶点 u 和 v 之间的没有相同边的两条简单通路。证明：在 G 里存在简单回路。

44. 证明：长度为 k 的简单回路的存在性是一个同构不变量，其中 k 是大于 2 的正整数。

45. 解释一下如何用定理 2 判定图是否连通。

46. 用练习 45 证明：图 8-40 里的图 G_1 是连通的而图 G_2 不是连通的。

47. 证明：简单图 G 是偶图，当且仅当 G 没有包含奇数条边的回路。

8.5 欧拉通路与哈密顿通路

8.5.1 引言

能否从一个顶点出发沿着图的边前进，恰好经过图的每条边一次并且回到这个顶点？同样，能否从一个顶点出发沿着图的边前进，恰好经过图的每个顶点一次并且回到这个顶点？虽然这两个问题有相似之处，但是对于所有的图来说，可以轻而易举地回答第一个关于是否具有欧拉回路的问题，却非常难以解决第二个关于是否具有哈密顿回路的问题。本节将研究这些问题并且讨论这些问题的难点。虽然两个问题在许多不同领域里都有实际应用，但是都来源于古老的智力题。下面将介绍这些古老的智力题以及现代的实际应用。

8.5.2 欧拉通路与欧拉回路

普鲁士的哥尼斯堡镇（现名加里宁格勒，属于俄罗斯共和国）被普雷格尔河支流分成四部分。这四部分包括河两岸、河中心岛以及两条支流之间的部分。在 18 世纪这四部分用七座桥连接起来，如图 8-47 所示。

镇上的人们在周日里穿过镇子进行长距离的散步。他们想弄明白是否可能从镇里某个位置出发不重复地经过所有桥并且返回出发点。

瑞士数学家列昂哈德·欧拉[⊖]解决了这个问题。他的解答发表在 1736 年，这也许是人们第一次使用图论。欧拉利用多重图来研究这个问题，其中用顶点表示四个部分，用边表示桥，如图 8-48所示。

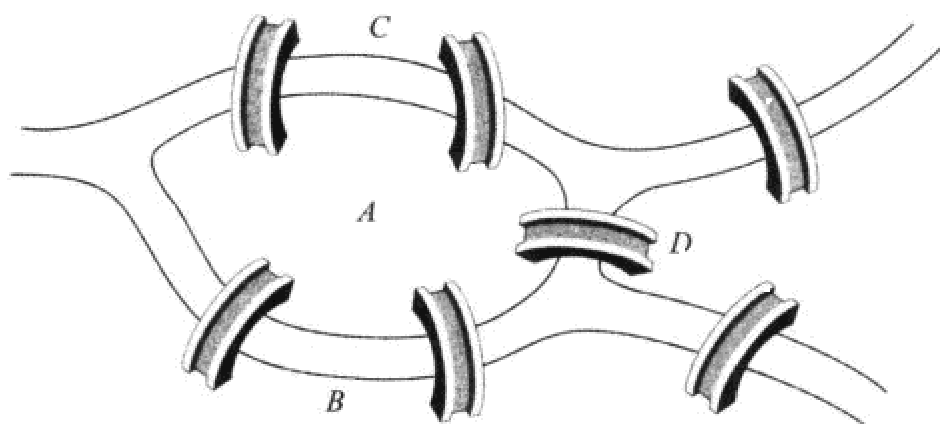


图 8-47 哥尼斯堡七桥

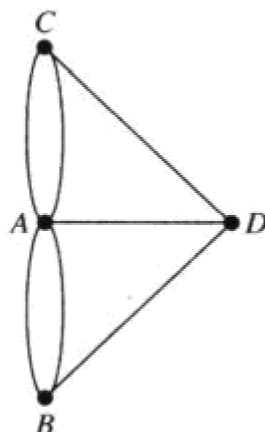


图 8-48 哥尼斯堡镇的多重图模型

不重复地经过每一座桥来旅行的问题可以利用这个模型来重新叙述。问题变成：在这个多重图里是否存在着包含每一条边的简单回路？

⊖ **列昂哈德·欧拉（Leonhard Euler，1707—1783）** 欧拉是瑞士巴塞尔附近一位加尔文教派牧师之子。他 13 岁进入巴塞尔大学，遵照父亲的愿望开始神学生涯。在大学里，欧拉受到著名的伯努利家族中的数学家约翰·伯努利的指导。他的兴趣和熟练的技能使他放弃神学研究而转向数学。欧拉 16 岁时取得了哲学硕士学位。1727 年彼得大帝邀请他加入圣彼得堡的科学院。1741 年他来到柏林科学院，在这里一直呆到 1766 年。然后他回到圣彼得堡，并在那里度过余生。

欧拉的成果多得令人难以置信。他在数学的许多领域都做出了贡献，包括数论、组合以及分析在诸如音乐和造船学这样的领域的应用等。他写的书籍和文章数量在 1100 部以上，另外还有非常多的去世前未能发表的著作，以至于在他去世之后，用了 47 年才发表完他的所有著作。他写文章非常快，在世时总有一大摞文章等待发表。柏林科学院总是先发表这一摞最顶上的文章，所以后来的结果常常先于它们所依赖或取代的结果而发表。欧拉有 13 个孩子，当有一两个孩子在他膝上玩耍时，他照样能够工作。在他生命的最后 17 年里，他完全失明了。但是，由于他神奇的记忆力，他的数学研究成果的推出并没有受到影响。他的全集的出版工作由瑞士自然科学协会负责，目前还在进行之中，预期将超过 75 卷。

定义1 图G里的欧拉回路是包含着G的每一条边的简单回路。图G里的欧拉通路是包含着G的每一条边的简单通路。

例1和例2解释了欧拉回路和欧拉通路的概念。

例1 在图8-49里，哪些无向图具有欧拉回路？在没有欧拉回路的那些图里，哪些具有欧拉通路？

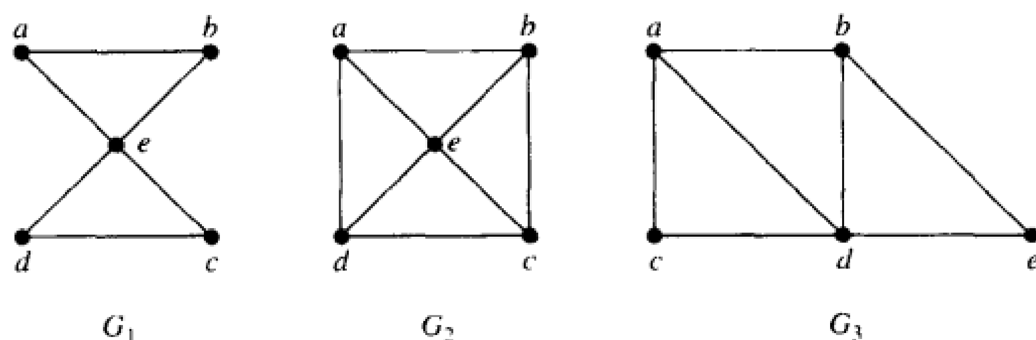


图8-49 无向图G_1，G_2和G_3

解 图G_1具有欧拉回路，例如a, e, c, d, e, b, a。G_2和G_3都没有欧拉回路(读者应当验证它)。但是G_3具有欧拉通路，即a, c, d, e, b, d, a, b。G_2没有欧拉通路(读者应当验证它)。■

例2 在图8-50里，哪些有向图具有欧拉回路？在没有欧拉回路的那些图里，哪些具有欧拉通路？

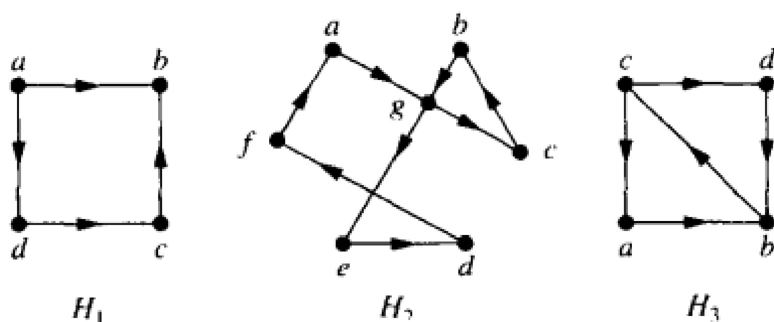


图8-50 有向图H_1，H_2和H_3

Extra Examples **解** 图H_2具有欧拉回路，例如$a, g, c, b, g, e, d, f, a$。$H_1$和$H_3$都没有欧拉回路(读者应当验证它)。$H_3$具有欧拉通路，即$c, a, b, c, d, b$，但是$H_1$没有欧拉通路(读者应当验证它)。■

欧拉回路和欧拉通路的充要条件 对判断多重图是否具有欧拉回路和欧拉通路来说，存在着简单的标准。欧拉在解决著名的哥尼斯堡七桥问题时发现了它们。假设在本节里讨论的所有图都具有有穷多个顶点和边。

若一个连通多重图具有欧拉回路，则它有什么性质呢？可以说明的是：每一个顶点都必有偶数条边。为此，首先注意一条欧拉回路从顶点a开始，接着是a关联的一条边，比方说$\{a, b\}$。边$\{a, b\}$为$\deg(a)$贡献1度。这条回路每次经过一个顶点就为该顶点的度贡献2度，这是因为这条回路经过关联该顶点的边进入，又经过另一条这样的边离开。最后，这条回路在它开始的地方结束，为$\deg(a)$贡献1度。因此$\deg(a)$必为偶数，这是因为当回路开始时它贡献1度，当回路结束时它贡献1度，每次经过a都贡献2度(如果它又经过了a)。除a外的其余顶点都有偶数度，这是因为每次回路经过一个顶点就为该顶点的度贡献2度。由此得出结论若连通图有欧拉回路，则每一个顶点必有偶数度。

欧拉回路的存在性的这个必要条件是否也是充分的？即若在连通多重图里所有顶点都有偶数度，则是否必有欧拉回路？这个问题可以通过构造来解决。

假设 G 是连通多重图而且 G 的每一个顶点都有偶数度。构造从 G 的任意顶点 a 开始的简单回路。设 $x_0=a$。首先任意地选择一条关联 a 的边 $\{x_0, x_1\}$。通过建立尽量长的简单通路 $\{x_0, x_1\}$，$\{x_1, x_2\}$，…，$\{x_{n-1}, x_n\}$来继续构造。例如，在图 8-51 的图 G 里，从 a 开始而且连续地选择边 $\{a, f\}$，$\{f, c\}$，$\{c, b\}$和$\{b, a\}$。

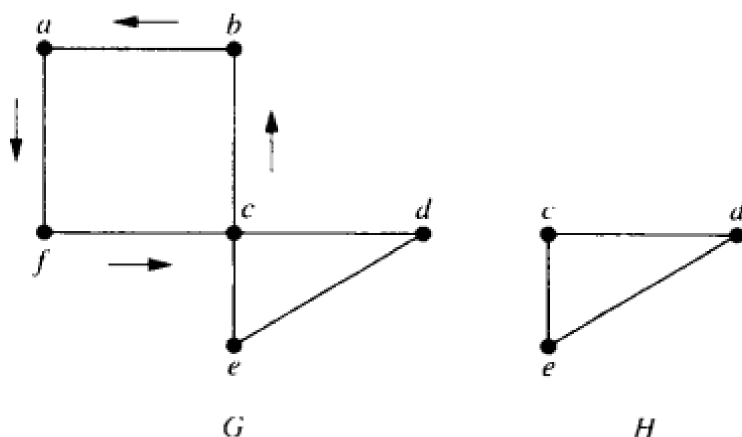


图 8-51　构造 G 里的欧拉回路

这样的通路必然结束，这是因为图的边数是有穷的。它在 a 上以形如$\{a, x\}$的边开始，而且在 a 上以形如$\{y, a\}$的边结束。这是因为通路每次经过一个偶数度顶点时，它只用 1 条边进入这个顶点，所以至少还剩下 1 条边让通路离开这个顶点。这条通路可能用完了所有的边，也可能没用完。

若所有的边都已经用完，则欧拉回路已经构造好了。否则，考虑通过从 G 里删除已经用过的边和不关联任何剩余的边的顶点这样所得到的子图 H。当从图 8-51 里的图删除回路 a，f，c，b，a 时，就得到标记为 H 的子图。

因为 G 是连通的，所以 H 与已经删除的回路至少具有 1 个公共顶点。设 w 是这样的顶点（此例中 c 是这个顶点）。

H 里的每一个顶点都有偶数度。（因为 G 里的所有顶点都有偶数度，对每个顶点来说，把与这个顶点关联的边成对地删除，以便形成 H。）注意 H 可能是不连通的。像在 G 里做过的那样，在 w 上开始，通过尽可能地选择边来构造 H 里的简单回路。这条回路必然在 w 上结束。例如，在图 8-51 里，c，d，e，c 是 H 里的回路。下一步通过把 H 里的回路与 G 里原来的回路拼接起来形成 G 里的回路（这是可行的，因为 w 是这个回路的顶点之一）。当在图 8-51 的图里这样做时，就得到回路 a，f，c，d，e，c，b，a。

继续进行这个过程，直到已经用完了所有的边为止（这个过程必然结束，这是因为图里只有有穷的边数）。这样就产生出欧拉回路。这样的构造说明，若连通多重图的顶点都有偶数度，则该图具有欧拉回路。

把这些结果总结成定理 1。

定理 1　连通多重图具有欧拉回路当且仅当它的每个顶点都有偶数度。

现在可以解决哥尼斯堡七桥问题了。因为图 8-48 所示的表示这些桥的多重图具有 4 个奇数度顶点，所以它没有欧拉回路。从给定点开始、恰好经过每座桥一次并返回开始点是无法实现的。

算法 1 给出了在定理 1 之前的讨论里所给的求欧拉回路的构造过程。（因为这个过程里的回路是任意地选择的，所以存在一些不确定性。这里不再通过更精确地说明过程的步骤来消除这些不确定性）。

算法 1 构造欧拉回路
procedure $Euler(G$：所有顶点有偶数度的连通多重图）
circuit：=在 G 里任选的顶点开始，连续地加入边所形成的回到该顶点的回路
H：=删除这条回路的边之后的 G
while H 还有边
begin
  subcircuit：=在既是 H 里的顶点也是 *circuit* 的边的端点处开始的 H 里的一条回路
  H：=删除 *subcircuit* 的边和所有孤立点之后的 H
  circuit：=在适当顶点上插入 *subcircuit* 之后的 *circuit*
end{*circuit* 是欧拉回路}

下一个例子说明如何利用欧拉通路和欧拉回路来解决一种类型的智力题。

例 3 有许多智力题要求用铅笔连续移动，不离开纸面并且不重复地画出图形。利用欧拉回路和欧拉通路来解决这样的智力题。例如，能否用这样的方法画出图 8-52 所示的穆罕默德短弯刀？其中图形是在同一个顶点上开始和结束。

解 可以解决这个问题，因为图 8-52 所示的图 G 具有欧拉回路。它具有这样的回路是因为它的所有顶点都有偶数度。用算法 1 来构造欧拉回路。首先形成回路 $a, b, d, c, b, e, i, f, e, a$。通过删除这条回路的边并且删除因此产生的孤立点，就得到子图 H。然后形成 H 里的回路 $d, g, h, j, i, h, k, g, f, d$。形成这条回路之后就用完了 G 里的所有边。在适当的地方拼接这条回路和第一条回路，就产生出欧拉回路 $a, b, d, g, h, j, i, h, k, g, f, d, c, b, e, i, f, e, a$。这条回路给出了铅笔不离开纸面并且不重复地画出弯刀的方法。■

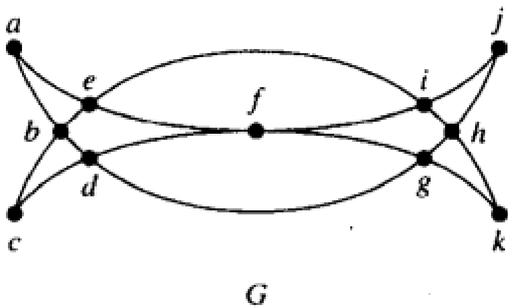


图 8-52 穆罕默德短弯刀

构造欧拉回路的另一个算法称为弗勒里算法，在本节末尾的练习里描述它。

现在说明，连通多重图具有欧拉通路（不是欧拉回路）当且仅当它恰有 2 个奇数度顶点。首先假设连通多重图有从 a 到 b 的欧拉通路，但不是欧拉回路。通路的第一条边为 a 的度贡献 1 度。通路每次经过 a 就为 a 的度贡献 2 度。通路的最后一条边为 b 的度贡献 1 度。通路每次经过 b 就为 b 的度贡献 2 度。所以 a 和 b 的度都是奇数。每一个其他顶点都具有偶数度，这是因为每当通路经过一个顶点时，就为这个顶点的度贡献 2 度。

现在反过来考虑。假设这个图恰有 2 个奇数度顶点，比方说 a 和 b。考虑由原来的图和边 $\{a, b\}$ 所组成的更大的图。这个更大的图的每一个顶点都有偶数度，所以具有欧拉回路。删除新边就产生原图的欧拉通路。定理 2 总结了这些结果。

定理 2 连通多重图具有欧拉通路但无欧拉回路当且仅当它恰有 2 个奇数度顶点。

例 4 图 8-53 所示的哪些图具有欧拉通路？

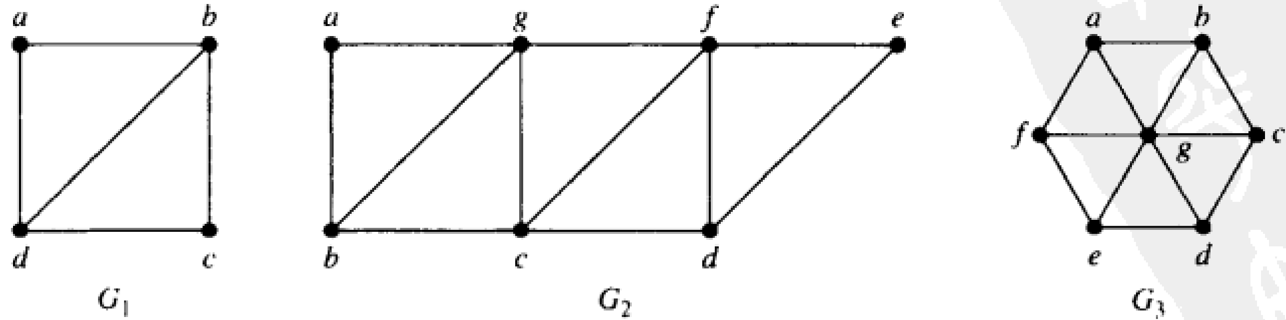


图 8-53 三个无向图

解 G_1 恰有 2 个奇数度顶点，即 b 和 d。因此它具有必须用 b 和 d 作为端点的欧拉通路。一条这样的欧拉通路是 d，a，b，c，d，b。同理，G_2 恰有 2 个奇数度顶点，即 b 和 d。因此它具有必须用 b 和 d 作为端点的欧拉通路。一条这样的欧拉通路是 b，a，g，f，e，d，c，g，b，c，f，d。G_3 没有欧拉通路，因为它具有 6 个奇数度顶点。 ■

回到 18 世纪的哥尼斯堡，是否有可能在镇里某点开始，旅行经过所有的桥，在镇里其他某点结束？通过判定表示哥尼斯堡七桥的多重图是否具有欧拉通路，就可以回答这个问题。因为这个多重图有四个奇数度顶点，没有欧拉通路，所以这样的旅行是不可能的。

有向图里欧拉通路和欧拉回路的充要条件，在本节末尾的练习里讨论。



可以用欧拉通路和欧拉回路解决许多实际问题。例如，许多应用都要求一条通路或回路，它要恰好一次地经过一个街区里的每条街道、一个交通网里的每条道路、一个高压输电网里的每个连接或者一个通信网络里的每个链接。求出适当的图模型里的欧拉通路或欧拉回路就可以解决这样的问题。例如，如果一个邮递员可以求出表示他所负责投递的街道的图中的欧拉通路，则这条通路就产生恰好一次经过每条街道的投递路线。这个问题称为*中国邮递员问题*，以纪念在 1962 年提出这个问题的中国科学家管梅谷。参看 [MiRo91] 了解关于不存在欧拉通路时中国邮递员问题的解的更多信息。

应用欧拉通路和欧拉回路的其他领域有：电路布线、网络组播和分子生物学，在分子生物学中欧拉通路用于 DNA 测序。

8.5.3 哈密顿通路与哈密顿回路



包含多重图每一条边恰好一次的通路和回路的存在性的充要条件已经得出。那么包含图的每一个顶点恰好一次的简单通路和回路的存在性的充要条件是否也能得出呢？

定义 2 在图 $G=(V, E)$ 里，若 $V=\{x_0, x_1, \cdots, x_{n-1}, x_n\}$ 并且对 $0\leqslant i<j\leqslant n$ 来说有 $x_i\neq x_j$，则通路 x_0，x_1，$\cdots$，x_{n-1}，x_n 称为*哈密顿通路*。在图 $G=(V, E)$ 里，若 x_0，x_1，$\cdots$，x_{n-1}，x_n 是哈密顿通路，则 x_0，x_1，$\cdots$，x_{n-1}，x_n，x_0（其中 $n>1$）称为*哈密顿回路*。

这个术语来自爱尔兰数学家威廉·罗万·哈密顿⊖爵士在 1857 年发明的智力题。哈密顿的智力题用到了木质十二面体（如图 8-54a 所示，十二面体有 12 个正五边形表面），十二面体每个



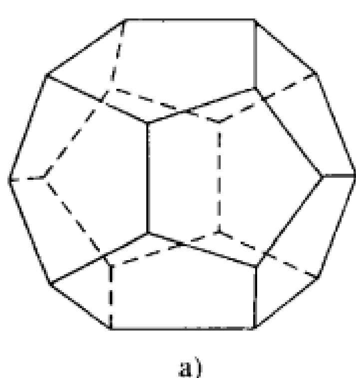
⊖ **威廉·罗万·哈密顿（William Rowan Hamilton，1805—1865）** 哈密顿这位爱尔兰最具名望的科学家 1805 年出生在都柏林。他的父亲是成功的律师，母亲来自以智力超群而闻名的家族，他本人是个神童。3 岁时他就是一名出色的读者并掌握了高等算术。因为聪明，他被送到身为著名语言学家的叔叔詹姆士那里生活。到 8 岁时，哈密顿学会了拉丁语、希腊语和希伯来语；到 10 岁时他又学会了意大利语和法语，并且开始学习东方语言，包括阿拉伯语、梵语和波斯语。在此期间他以懂得当时的所有语言而自豪。17 岁时他不再学习新的语言，但是已经掌握了数学分析和许多数学天文学知识。他开始了在光学上的开创性工作，发现了拉普拉斯的天体力学著作中的重大错误。在 18 岁进入都柏林三一学院之前，哈密顿一直接受私人教育。在三一学院里他在科学和古典文学上都表现超群。在获得学位之前，他就因为过人的才华，从多位著名天文学家参与的竞争中脱颖而出，被任命为爱尔兰皇家天文学家。他终身担任这个职位，在都柏林郊外的邓辛天文台生活和工作。哈密顿在光学、抽象代数和动力学领域做出了重要贡献。哈密顿发明了称为四元数的代数对象来作为非交换系统的例子。当他沿都柏林的运河散步时，发现了四元数相乘的适当方式。狂喜之下，他把公式刻在了跨越运河的石桥上，今日该地立匾为记。随后哈密顿一直沉迷在四元数里，努力把它们应用到数学的其他领域，而不再转向新的研究领域。

1857 年，哈密顿在自己非交换代数的工作的基础上发明了"艾口西安游戏"。他把这个想法以 25 镑的价格出售给游戏和智力题的经销商。（游戏的销路一直不好，事实证明这是经销商的一次失败的投资。）本节所描述的智力题"旅行者十二面体"，又称"周游世界"，就是该游戏的变种。

哈密顿在 1833 年第三次结婚，但是他的婚姻很不幸，因为他的妻子是半残疾人，无法处理他的家务。他在生命的最后 20 年里，过着酗酒和隐居的生活。他 1865 年死于痛风，留下大量包含未发表研究结果的文稿。在这些文稿里，混杂着大量晚餐碟子，许多碟子里还有已脱水的吃剩的排骨。

顶点上的钉子，以及细线。十二面体的 20 个顶点用世界上的不同城市作标记。智力题要求从一个城市开始，沿十二面体的边旅行，访问其他 19 个城市每个恰好一次，回到第一个城市结束。旅行经过的回路用钉子和细线来标记。

因为作者不可能向每位读者提供带钉子和细线的木质十二面体，所以考虑一个等价的问题：图 8-54b 中的图是否具有恰好经过每个顶点一次的回路？它就是对原题的解，因为该图同构于包含十二面体顶点和边的图。图 8-55 是哈密顿智力题的解。


a)

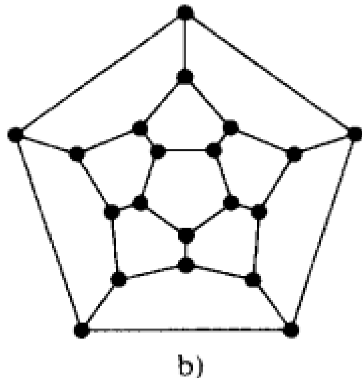

b)

图 8-54 哈密顿的"周游世界"智力题

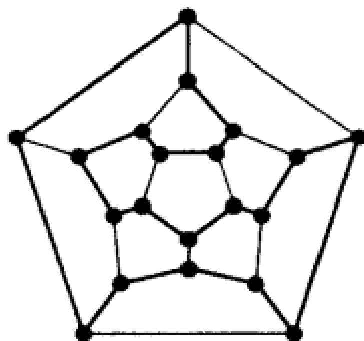
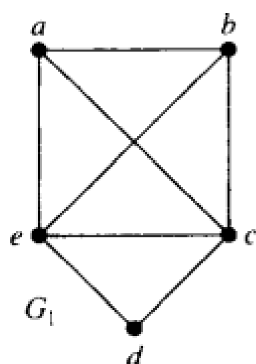
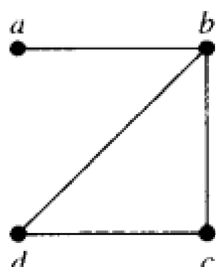


图 8-55 "周游世界"智力题的解

例 5 在图 8-56 里，哪些简单图具有哈密顿回路？或者没有哈密顿回路但是有哈密顿通路？





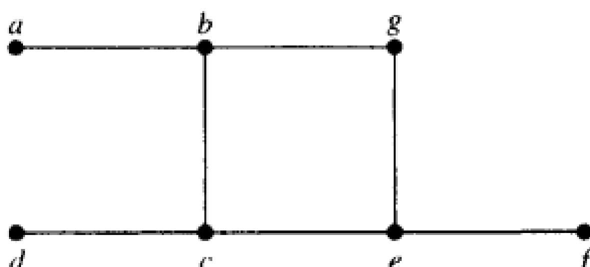
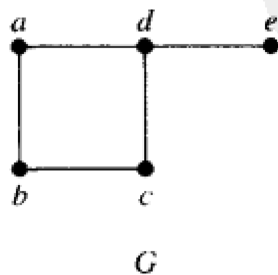


图 8-56 三个简单图

解 G_1 有哈密顿回路：a，b，c，d，e，a。G_2 没有哈密顿回路(从包含每一个顶点的任何回路必然两次包含边$\{a, b\}$可以看出这一点)，但是 G_2 确实有哈密顿通路，即 a，b，c，d。G_3 既无哈密顿回路也无哈密顿通路，这是因为包含所有顶点的任何通路都必然多次包含边$\{a, b\}$，$\{e, f\}$和$\{e, d\}$其中之一。 ■

是否存在简单方式来判定一个图有无哈密顿回路或哈密顿通路？初看起来，似乎应当有判定这一点的简单方式，因为存在简单方式来回答一个图有无欧拉回路这样的相似问题。令人吃惊的是，没有已知的简单的充要条件来判定哈密顿回路的存在性。不过，已经有许多定理对哈密顿回路的存在性给出了充分条件。另外，某些性质可以用来证明一个图没有哈密顿回路。例如，带有 1 度顶点的图没有哈密顿回路，因为在哈密顿回路里每个顶点都关联着回路里的两条边。另外，若图中有 2 度顶点，则关联这个顶点的两条边属于任意一条哈密顿回路。此外注意，当构造哈密顿回路而且该回路经过某一个顶点时，除了回路所用到的两条边以外，这个顶点所关联的其他所有边不用再考虑。而且，哈密顿回路里不能包含更小的回路。



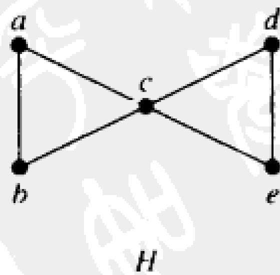


图 8-57 两个没有哈密顿回路的图

例 6 证明图 8-57 里的图都没有哈密顿回路。

解 G 里没有哈密顿回路，因为 G 有 1 度顶点，

即 e。现在考虑 H。因为顶点 a，b，d 和 e 的度都为 2，所以这些顶点关联的每一条边都必然属于任意一条哈密顿回路。现在容易看出 H 里不存在哈密顿回路，因为任何这样的哈密顿回路都不得不包含 4 条关联 c 的边，这是不可能的。■

例 7 证明：每当 $n\geqslant 3$ 时，K_n 就有哈密顿回路。

解 从 K_n 里的任意一个顶点开始来形成哈密顿回路。以所选择的任意顺序来访问顶点，只要求通路在同一个顶点开始和结束，而且对其他每个顶点恰好访问一次。这样做是可能的，因为在 K_n 里任意两个顶点之间都有边。■

虽然还未发现任何有用的关于哈密顿回路存在性的充要条件，但很多充分条件已经找到了。注意到一个图的边越多，这个图就越可能有哈密顿回路。另外，加入边(而不是顶点)到已经有哈密顿回路的图中就产生有相同哈密顿回路的图。因此，当加入边到图中时，特别是当确保给每个顶点都加入边时，这个图存在哈密顿回路的可能性就更大了。这里叙述充分条件中最重要的两个。这些条件是加布里尔・A・狄拉克(Gabriel A. Dirac)在 1952 年和奥伊斯坦・奥尔(Oystein Ore)在 1960 年发现的。

定理 3 狄拉克定理。如果 G 是带 n 个顶点的连通简单图，其中 $n\geqslant 3$，并且 G 中每个顶点的度都至少为 $n/2$，则 G 有哈密顿回路。

定理 4 奥尔定理。如果 G 是带 n 个顶点的连通简单图，其中 $n\geqslant 3$，并且对于 G 中每一对不相邻的顶点 u 和 v 来说，都有 $\deg(u)+\deg(v)\geqslant n$，则 G 有哈密顿回路。

本节末的练习 61 粗略介绍了奥尔定理的证明。狄拉克定理可以作为奥尔定理的推论来证明，因为狄拉克定理的条件蕴含了奥尔定理的条件。

奥尔定理和狄拉克定理都给出了连通简单图有哈密顿回路的充分条件。但是，这些定理没有给出哈密顿回路存在性的必要条件。例如，图 C_5 有哈密顿回路，但既不满足奥尔定理的假设也不满足狄拉克定理的假设，读者可以验证这一点。

已知的最好的求一个图里的哈密顿回路或判定这样的回路不存在的算法具有指数的最坏情形时间复杂性(相对于图的顶点数来说)。找到具有多项式最坏情形时间复杂性的解决算法是 NP 复杂的，它的发现将具有重大意义，因为这意味着其他许多难题都会迎刃而解。

可以用哈密顿通路和哈密顿回路来解决实际问题。例如，许多应用都要求一条通路或回路，它要恰好一次地访问一个城市里的每个路口、一个设备网格里的每个管道交汇处或者一个通信网络里的每个节点。求出适当的图模型里的哈密顿通路或哈密顿回路就可以解决这样的问题。著名的旅行商问题要求一个旅行商为了访问一组城市所应当选取的最短路线。这个问题可归结为求完全图的哈密顿回路，使得这个回路的边的权的总和尽可能地小。

现在给出哈密顿回路在编码上的一种不太明显的应用。

例 8 格雷码 旋转的指针的位置可以表示成数字的形式。一种方式是把圆周等分成 2^n 段弧并且用长度为 n 的位串给每段弧赋值。图 8-58 里显示出了用长度为 3 的位串来这样做的两种方式。

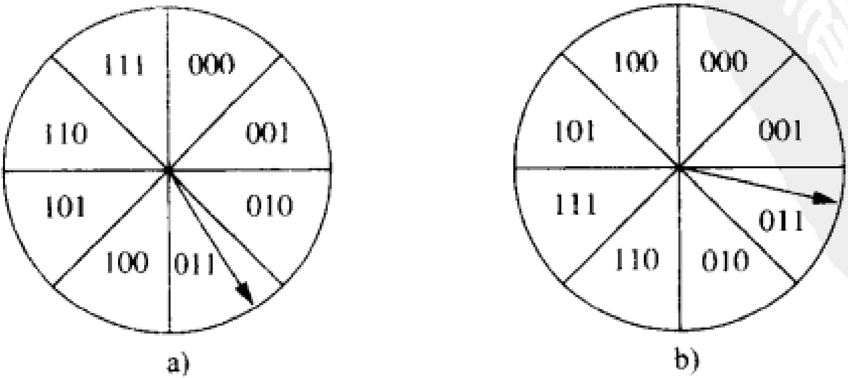


图 8-58 把指针位置转换成数字形式

用 n 个接触点来确定指针位置的数字表示。每个接触点用来读出位置的数字表示中的一位。图 8-59 针对图 8-58 里的两种赋值对此做了解释。

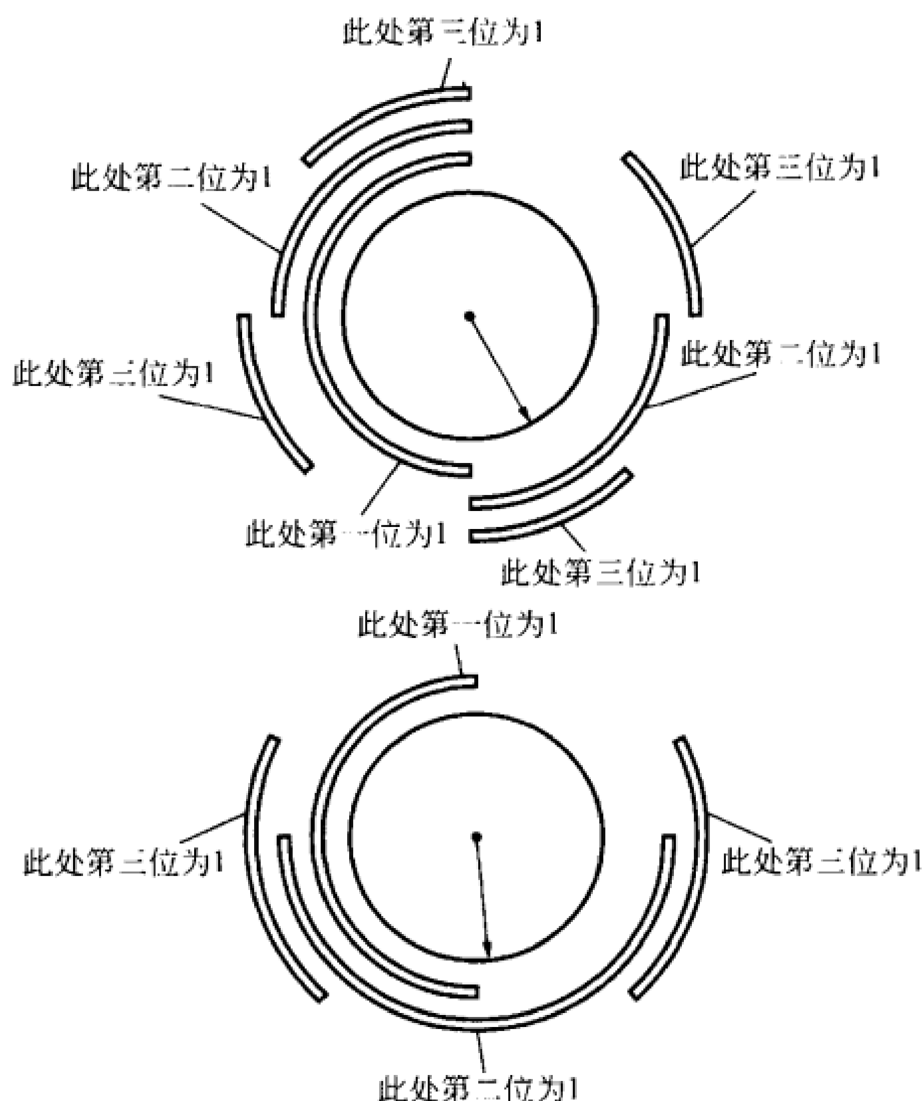


图 8-59 指针位置的数字表示

当指针靠近两段弧的边界时，在读出指针位置时可能发生错误。这可能引起读出的位串里有一个大的错误。例如，在图 8-58a 的编码方案里，若在确定指针位置的过程里发生了一个小的错误，则读出的位串是 100 而不是 011。所有三位都是错的！为了把在确定指针位置的过程里的错误的影响降到最低，用位串对 2^n 段弧的赋值，应当使相邻的弧所表示的位串只相差一位。在图 8-58b 里的编码方案的情况恰好就是这样的。在确定指针位置的过程里的一个错误使给出的位串为 010 而不为 011。只有一位是错的。

格雷码是圆周的弧的一种标记，使得相邻的弧具有恰好相差一位的位串标记。在图 8-58b 里的赋值是一个格雷码。可以这样找出格雷码：以下述的方式列出所有长度为 n 的位串，使得每一个串与前一个位串恰好相差一位，而且最后一个串与第一个位串恰好相差一位。可以用 n 立方体 Q_n 来为这个问题建模。解决这个问题所需要的是 Q_n 里的一条哈密顿回路。这样的哈密顿回路容易求出。例如，Q_3 的一条哈密顿回路显示在图 8-60 里。这条哈密顿回路所产生的前后恰好相差一位的位串序列是 000，001，011，010，110，111，101，100。

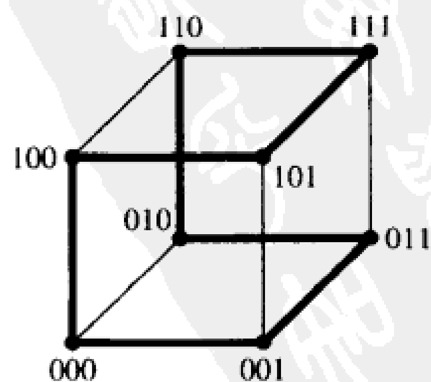
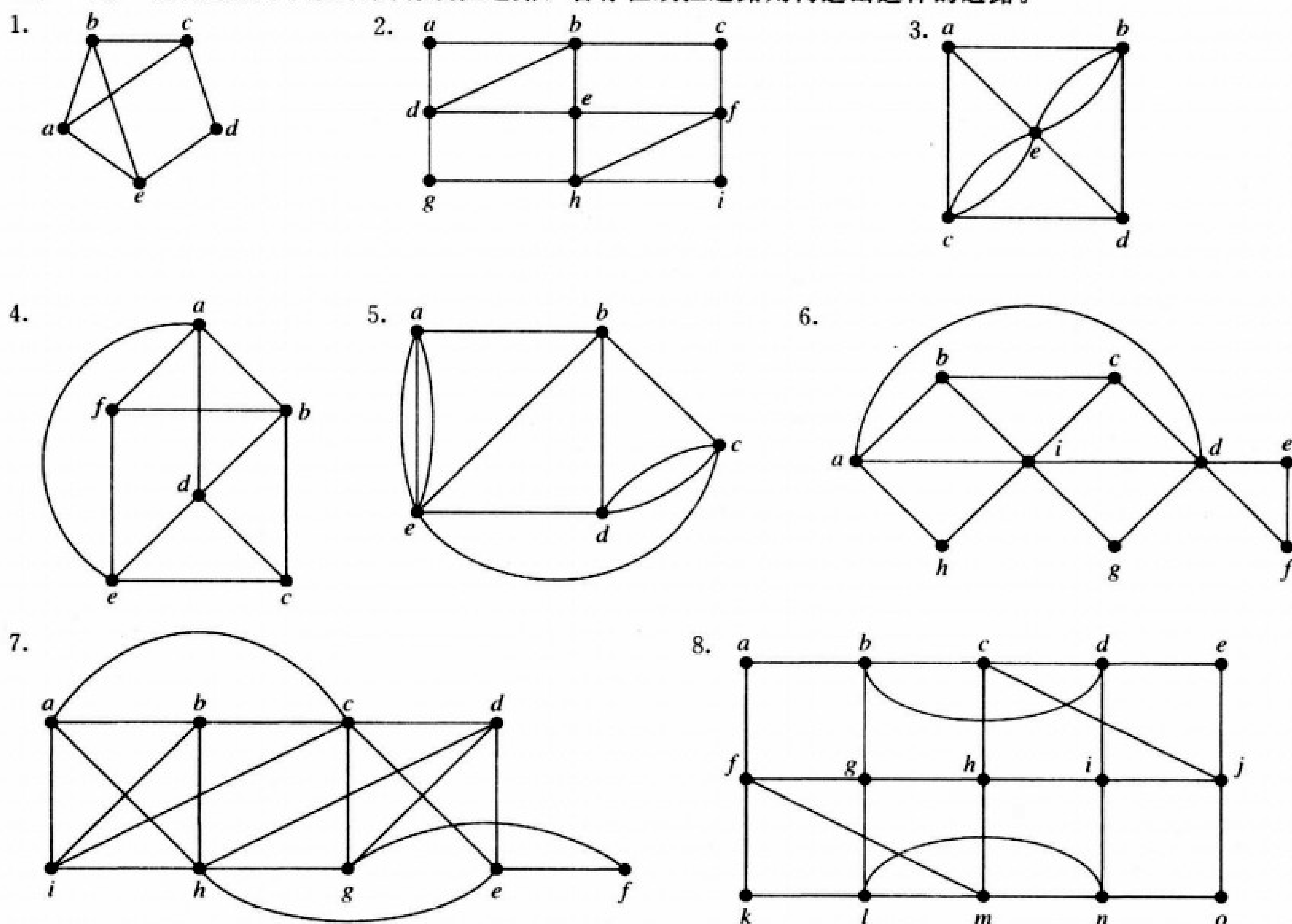


图 8-60 Q_3 的哈密顿回路

格雷码是以弗兰克·格雷的名字来命名的，20 世纪 40 年代，格雷在贝尔实验室研究如何把数字信号传送过程中错误的影响降到最低时发明了它们。 ■

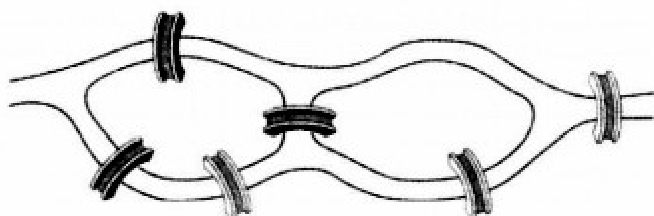
练习

在练习 1～8 里，判定给定的图是否具有欧拉回路。当存在欧拉回路时构造出这样的回路。如果不存在欧拉回路，就确定这个图是否具有欧拉通路，若存在欧拉通路则构造出这样的通路。



9. 在加里宁格勒(哥尼斯堡现名)，除了 18 世纪的七座桥之外，还有另外两座桥。这些新桥分别连接区域 B 和 C 以及区域 B 和 D。是否有人能够经过加里宁格勒的九座桥恰好一次并且回到出发点？

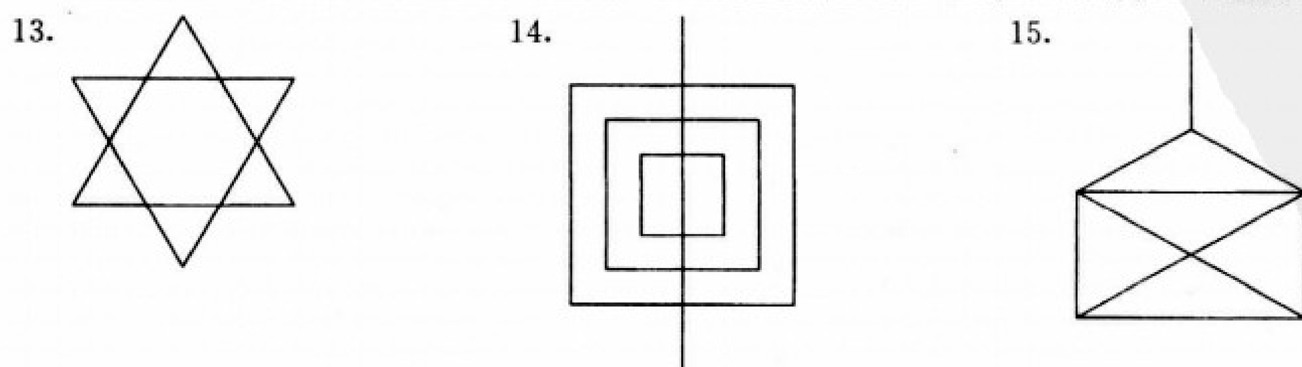
10. 是否有人能够经过下图所示的所有桥恰好一次并且回到出发点？



11. 何时可以画出一个城市里街道的中心线而不重复经过街道(假设所有街道都是双向街道)？

12. 设计一个与算法 1 相似的过程，它能够在多重图里构造欧拉通路。

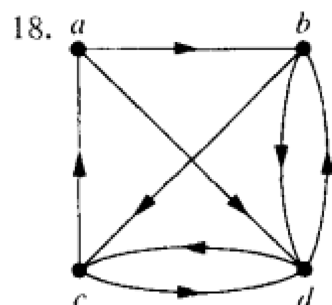
在练习 13～15 里，判定是否可以用一只铅笔连续移动，不离开纸面并且不重复地画出所示的图形。

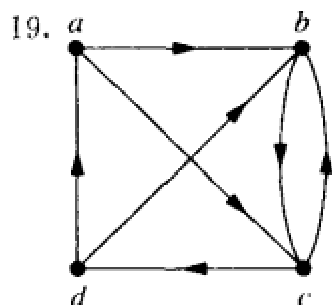


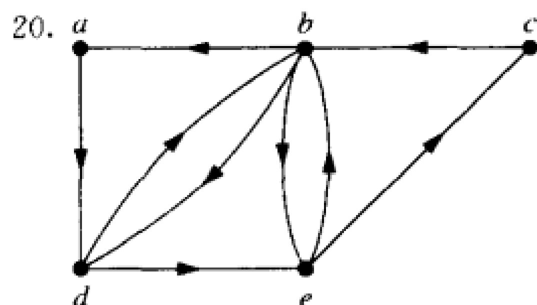
*16. 证明：不带有孤立点的有向多重图具有欧拉回路，当且仅当该图是弱连通的并且每个顶点的入度与出度都相等。

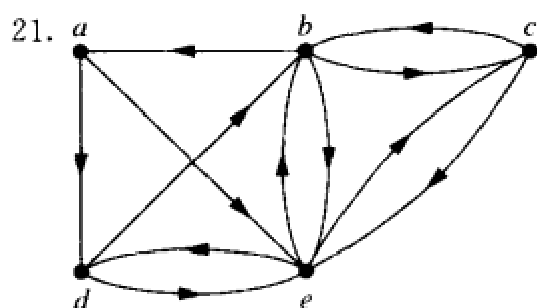
*17. 证明：不带有孤立点的有向多重图具有欧拉通路而没有欧拉回路，当且仅当该图是弱连通的并且存在两个顶点，一个顶点的入度比出度大 1 而另外一个顶点的出度比入度大 1，其余每个顶点的入度与出度都相等。

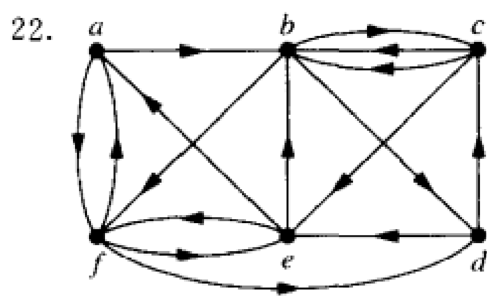
在练习 18～23 里，确定所示的有向图是否具有欧拉回路。若存在欧拉回路，则构造出一条欧拉回路。如果不存在欧拉回路，就确定这个有向图是否具有欧拉通路。若存在欧拉通路，则构造出一条欧拉通路。

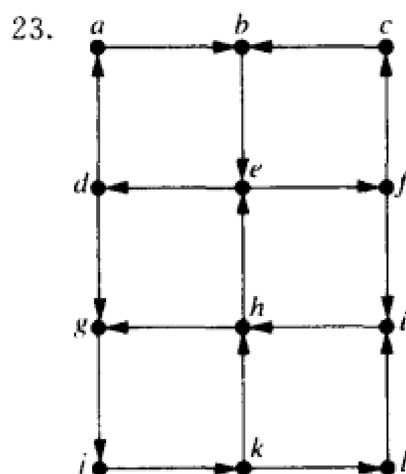
18.


19.


20.


21.


22.


23.


*24. 设计一个算法，它构造有向图里的欧拉回路。

25. 设计一个算法，它构造有向图里的欧拉通路。

26. 对哪些 n 值来说，下列图具有欧拉回路？

a) K_n　　b) C_n　　c) W_n　　d) Q_n

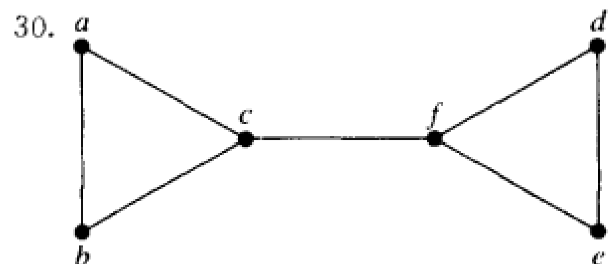
27. 对哪些 n 值来说，练习 26 里的图具有欧拉通路而没有欧拉回路？

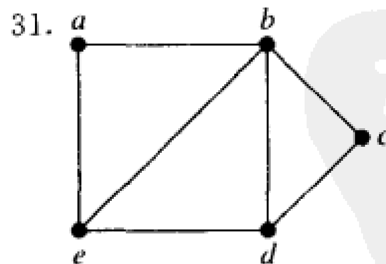
28. 对哪些 m 和 n 值来说，完全偶图 $K_{m,n}$ 具有

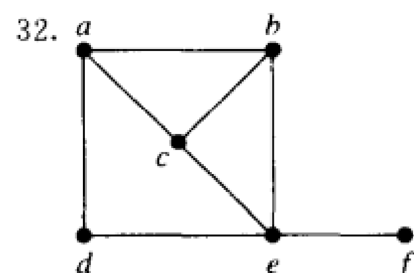
a) 欧拉回路？　　b) 欧拉通路？

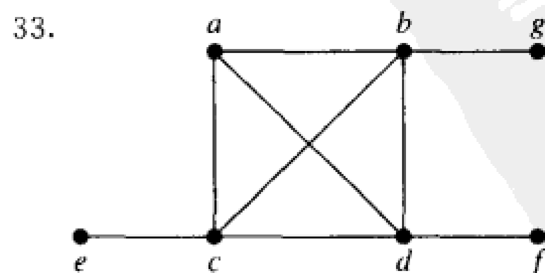
29. 当不重复任何部分地画出练习 1～7 里的每个图时，求出铅笔必须离开纸面的最少次数。

在练习 30～36 里，确定所给的图是否具有哈密顿回路。若有哈密顿回路，则求出这样一条回路。若没有哈密顿回路，则论证为什么这样的回路不存在。

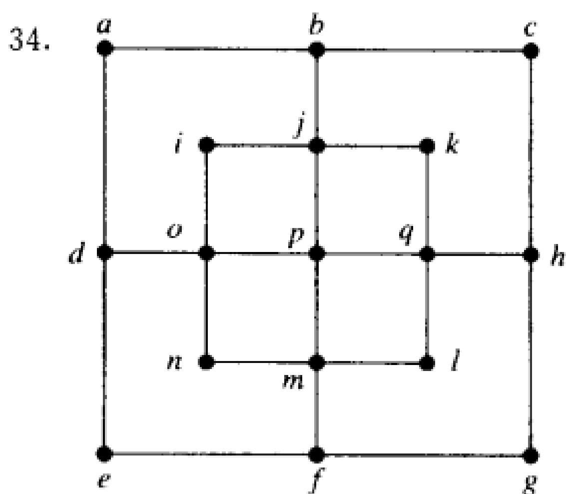
30.


31.


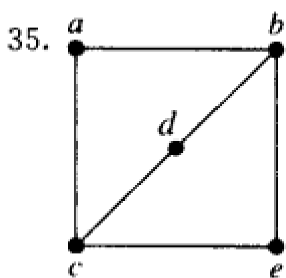
32.


33.


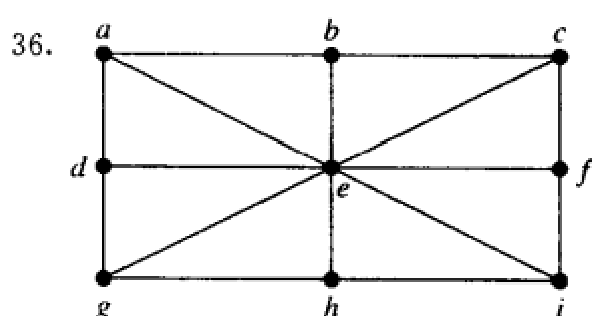
34.



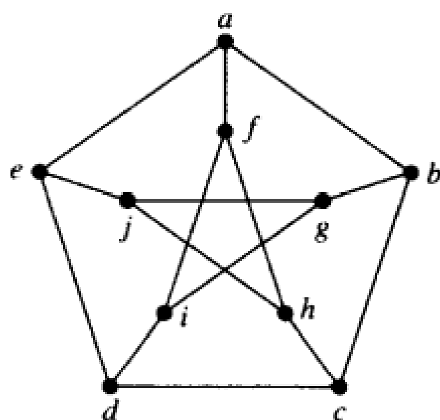
35.



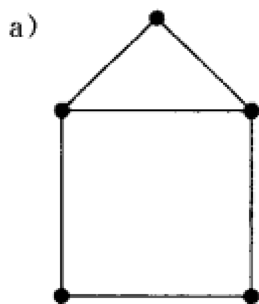
36.

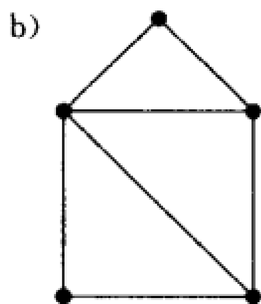


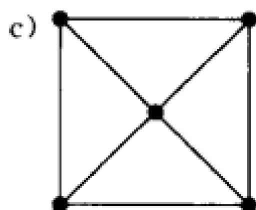
37. 练习 30 的图有哈密顿通路吗？如果有，找到它。如果没有，给出理由证明不存在这样的通路。
38. 练习 31 的图有哈密顿通路吗？如果有，找到它。如果没有，给出理由证明不存在这样的通路。
39. 练习 32 的图有哈密顿通路吗？如果有，找到它。如果没有，给出理由证明不存在这样的通路。
40. 练习 33 里的图是否具有哈密顿通路？若有哈密顿通路，则求出这样一条通路。若没有哈密顿通路，则给出论证来说明这样的通路为什么不存在。
*41. 练习 34 里的图是否具有哈密顿通路？若有哈密顿通路，则求出这样一条通路。若没有哈密顿通路，则论证为什么这样的通路不存在。
42. 练习 35 里的图是否具有哈密顿通路？若有哈密顿通路，则求出这样一条通路。若没有哈密顿通路，则论证为什么这样的通路不存在。
43. 练习 36 的图有哈密顿通路吗？如果有，找到它。如果没有，给出理由证明不存在这样的通路。
44. 对哪些 n 值来说，练习 26 里的图具有哈密顿回路？
45. 对哪些 m 和 n 值来说，完全偶图 $K_{m,n}$ 具有哈密顿回路？
*46. 证明：右图所示的彼得森[⊖]图没有哈密顿回路，但是删除顶点 v 和所有与 v 关联的边，所获得的子图却有哈密顿回路。

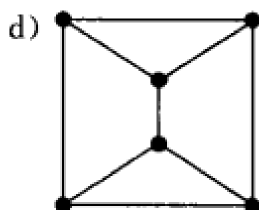


47. 对于下列每个图确定：(i)能否用狄拉克定理来证明这个图有哈密顿回路；(ii)能否用奥尔定理来证明这个图有哈密顿回路；(iii)这个图是否具有哈密顿回路。

a) 

b) 

c) 

d) 



⊖ **朱理乌斯·彼得·克里西安·彼得森(Julius Peter Christian Petersen，1839—1910)** 彼得森出生在丹麦的索镇。其父是染匠。在 1854 年父母再也负担不起他的学费，所以他到叔叔的杂货店当学徒。这位叔叔死的时候给彼得森留下足够的钱让他回到学校。毕业后他在哥本哈根工业学校开始学习工程，随后决定专攻数学。1858 年他出版了第一部教科书，一本关于对数的书。当继承的遗产用完之后，他不得不靠教书来谋生。从 1859 年直到 1871 年彼得森在哥本哈根的一所私立贵族高中教书。他一边教书一边继续研究，在 1862 年进入哥本哈根大学。1862 年他与劳拉·伯特森结婚；他们有三个孩子，两男一女。

1866 年彼得森从哥本哈根大学获得数学学位，并且在 1871 年从该校最终获得博士学位。得到博士学位后，他在工业与军事高等专科学校任教。1887 年他被任命为哥本哈根大学的教授。在丹麦，彼得森因著述有大量的高中和大学教科书而知名。其中《解决几何构造问题的方法和理论》被译成 8 种文字，英文版上次重印是在 1960 年，而法文版上次重印在 1990 年，距离初版日期超过一个世纪。

彼得森的研究领域很广泛，包括代数学、分析学、密码学、几何学、力学、数理经济学以及数论。他对图论的贡献，包括有关正则图的结果，最为著名。他以叙述的清晰性、解决问题的技巧性、独创性、幽默感、精力充沛以及擅长教学而闻名。彼得森不愿意读其他数学家的著作。所以，他经常发现别人已经证明过的结果，而常常陷于尴尬之中。不过当其他数学家不读他的作品时他却常常发怒！

彼得森去世的消息曾登在哥本哈根报纸的头版。当时一家报纸把他描述成科学界的汉斯·克里西安·安徒生——在学术世界里做出巨大贡献的人民之子。

48. 能否找到一个带有 n 个顶点的简单图($n \geqslant 3$)，这个图没有哈密顿回路，但是这个图中每个顶点的度至少是$(n-1)/2$?

*49. 证明：每当 n 是正整数时，就存在 n 阶格雷码，或者等价地证明：$n>1$ 的 n 立方体 Q_n 总是具有哈密顿回路。(提示：用数学归纳法。证明如何从 $n-1$ 阶格雷码产生 n 阶格雷码。)

构造欧拉回路的弗留利算法是从连通多重图的任意一个顶点开始，连续地选择边来形成一条回路。一旦选择了一条边，就删除这条边。连续地选择边，使得每条边从上一条边结束的地方开始，而且使得这条边不是一条割边，除非别无选择。

50. 用弗留利算法找出在图 8-51 中图 G 里的欧拉回路。

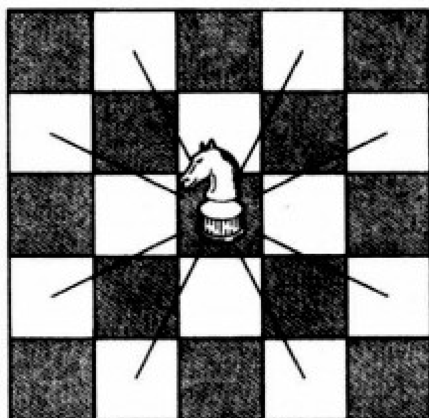
*51. 用伪代码表达弗留利算法。

**52. 证明：弗留利算法总是产生一条欧拉回路。

*53. 给出变化的弗留利算法来产生欧拉通路。

54. 一个诊断消息可以在计算机网络上发出，以便在所有连接和所有设备上执行测试。为了测试所有的连接，应当使用什么种类的通路？为了测试所有的设备呢？

55. 证明：带有奇数个顶点的偶图没有哈密顿回路。



马在国际象棋中是这样一种棋子，它的移动可以是水平两格加垂直一格，或者是水平一格加垂直两格。即在(x, y)格子的马可以移动到 8 个格子$(x\pm2, y\pm1)$，$(x\pm1, y\pm2)$中的任何一个，只要这些格子是在棋盘上，如右图所示。

马的周游是马的合法移动的序列，马在某个格子上开始而且访问每个格子恰好一次。若存在一种合法移动，它让马从周游的最后一个格子回到周游开始的地方，则马的周游称为*重返的*。可以用图为马的周游建模，棋盘上每个格子都用一个顶点来表示，若马可以在两个顶点所表示的格子之间合法地移动，则用一条边连接这两个顶点。

56. 画出表示马在 3×3 棋盘上的合法移动的图。

57. 画出表示马在 3×4 棋盘上的合法移动的图。

58. a)证明：求马在 $m\times n$ 棋盘上的周游等价于求表示马在该棋盘上合法移动的图的哈密顿通路。

    b)证明：求马在 $m\times n$ 棋盘上的重返的周游等价于求所对应的图上的哈密顿回路。

*59. 证明：存在马在 3×4 棋盘上的周游。

*60. 证明：不存在马在 3×3 棋盘上的周游。

*61. 证明：不存在马在 4×4 棋盘上的周游。

62. 证明：每当 m 和 n 都是正整数时，表示马在$m\times n$ 棋盘上的合法移动的图就是偶图。

63. 证明：当 m 和 n 都是奇数时，不存在马在$m\times n$ 棋盘上的重返的周游(提示：利用练习 51、练习 54b 和练习 58)。

*64. 证明：存在马在 8×8 棋盘上的周游。(提示：你可以用沃恩斯道夫于 1823 年发明的下列方法来构造马的周游。从任意格子里开始，然后总是移动到与最少数目的没有用过的格子连接的一个格子。虽然这个方法不能总是产生马的周游，但是它确实很有效。)

65. 本练习粗略介绍奥尔定理的证明。假设 G 是带有 n 个顶点的简单图，$n\geqslant3$，并且每当 x 和 y 是 G 中不相邻的顶点时，$\deg(x)+\deg(y)\geqslant n$。奥尔定理称在这些条件下 G 有哈密顿回路。

    a)证明：如果 G 没有哈密顿回路，则存在另一个带有与 G 相同顶点的图 H，可以这样来构造 H：加入边到 G 使得再加入一条边就产生 H 中的哈密顿回路。(提示：依次在 G 的每个顶点处加入不产生哈密顿回路的尽可能多的边。)

    b)证明：在 H 中存在哈密顿通路。

    c)设 v_1，v_2，…，v_n 是 H 中的哈密顿通路。证明：$\deg(v_1)+\deg(v_n)\geqslant n$ 并且至多存在 $\deg(v_1)$ 个顶点不与 v_n 相邻(包括 v_n 在内)。

d)设 S 是与哈密顿通路上 v_1 相邻的每个顶点前面的顶点的集合。证明 S 包含 $\deg(v_1)$ 个顶点并且 $v_n \notin S$。

e)证明：S 包含与 v_n 相邻的顶点 v_k。这蕴含着存在连接 v_1 与 v_{k+1} 和 v_k 与 v_n 的边。

f)证明：(e)蕴含着 v_1，v_2，…，v_{k-1}，v_k，v_n，v_{n-1}，…，v_{k+1}，v_1 是 G 中的哈密顿回路。从这个矛盾得出奥尔定理成立。

8.6 最短通路问题

8.6.1 引言

许多问题可以用边上赋权的图来建模。作为说明，考虑一下航线系统如何建模。如果用顶点表示城市，用边表示航班。给边赋上城市之间的距离，就可以为涉及距离的问题建模；给边赋上飞行时间，就可以为涉及飞行时间的问题建模；给边赋上票价，就可以为涉及票价的问题建模。图 8-61 显示了给一个图的边赋权的三种不同方式，分别表示距离、飞行时间和票价。

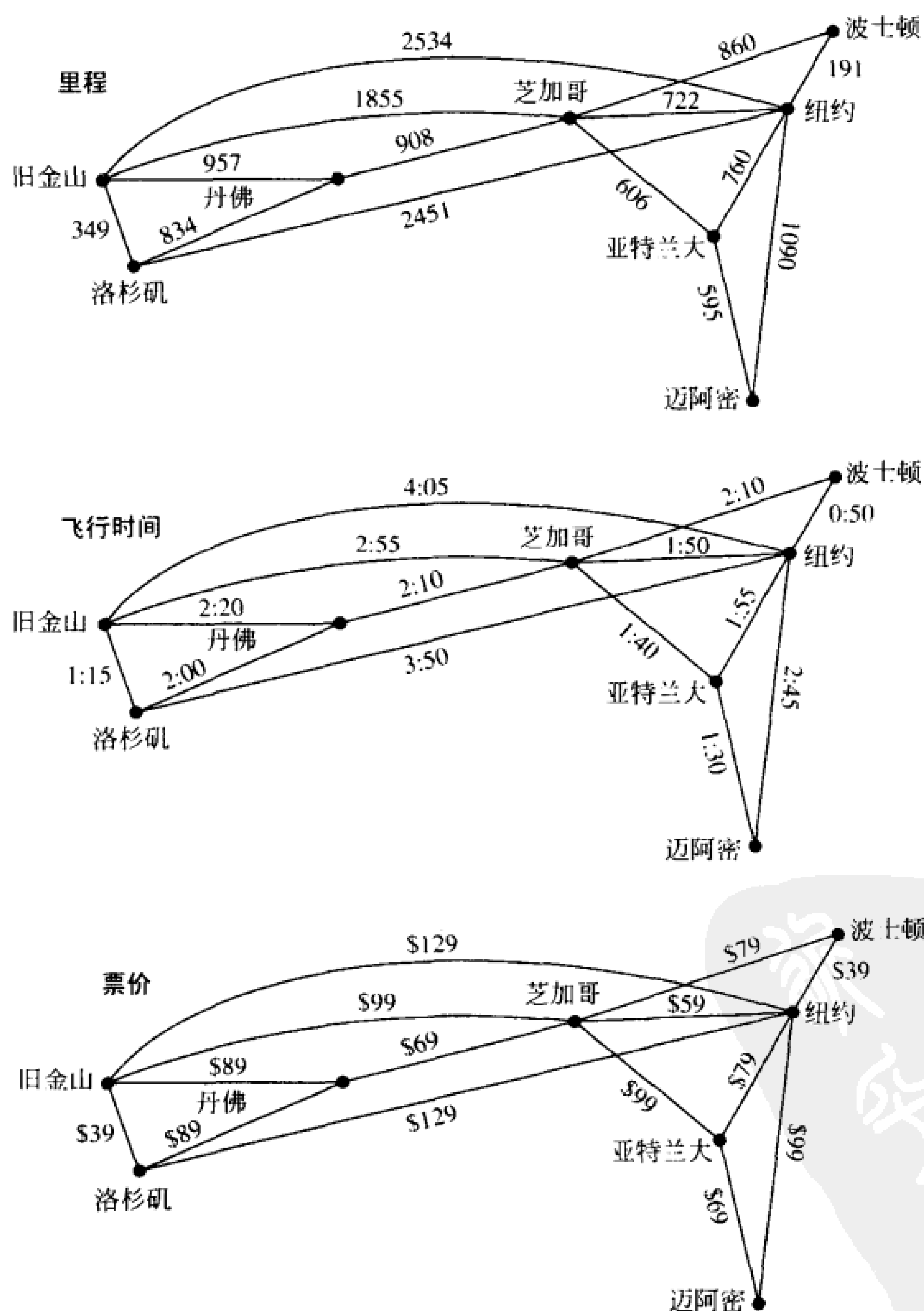


图 8-61　为航线系统建模的带权图

给每条边赋上一个数的图称为带权图。带权图用来为计算机网络建模。通信成本(比如租用电话线的月租费)、计算机在这些线路上的响应时间或计算机之间的距离等都可以用带权图来研究。图 8-62 显示一些带权图，它们表示给计算机网络的图的边赋权的三种方式，分别对应于成本、响应时间和距离。

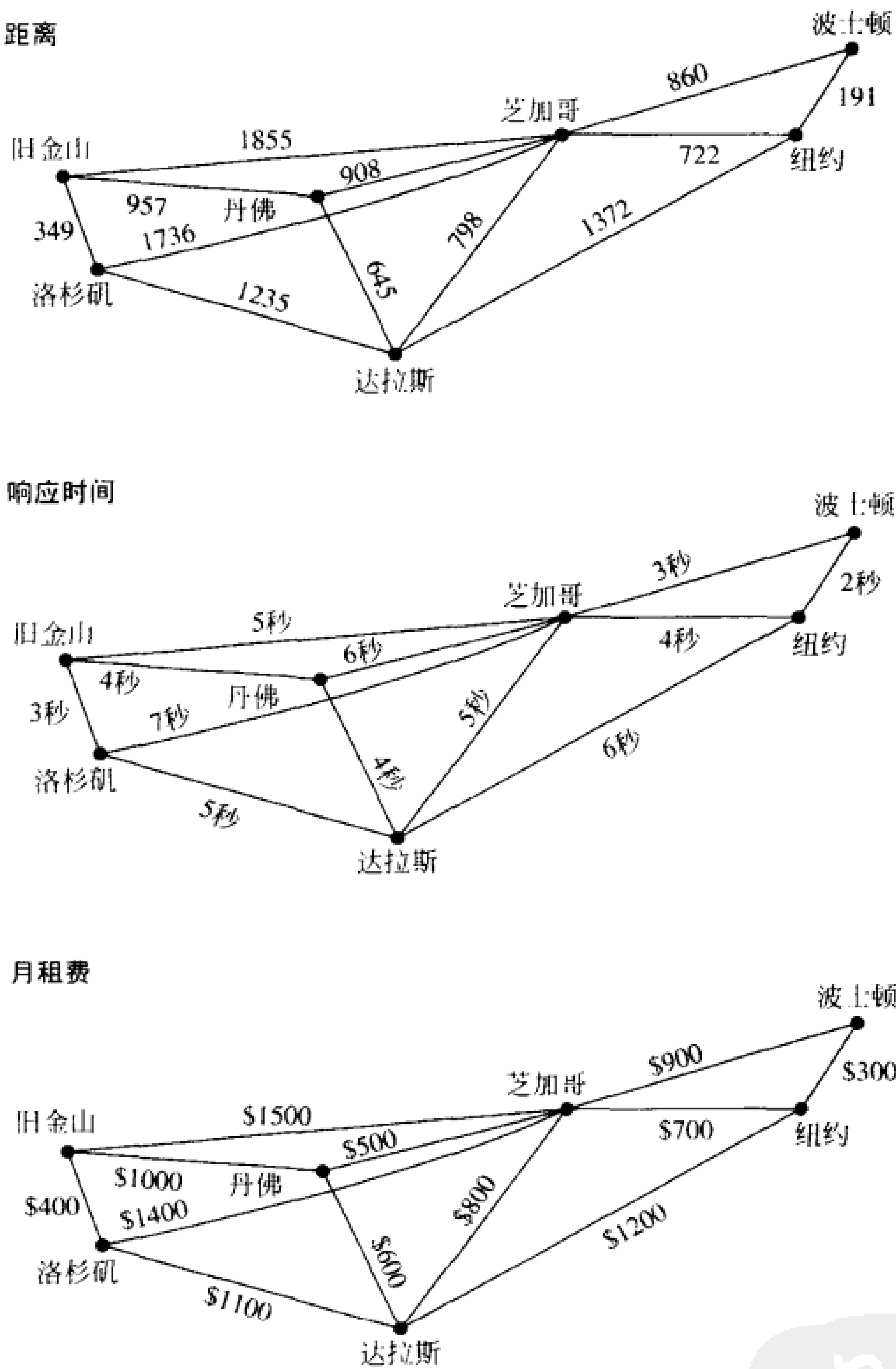


图 8-62 为计算机网络建模的带权图

与带权图有关的几种类型的问题出现得很多。确定网络里两个顶点之间长度最短的通路就是一个这样的问题。说得更具体些，设带权图里一条通路的长度是这条通路上各边的权的总和。(读者应当注意，对术语长度的这种用法，与表示不带权的图的通路里边数的长度的用法，这两者是不同的。)问题是：什么是最短通路，即什么是在两个给定顶点之间长度最短的通路？例如，在图 8-61 所示带权图表示的航线系统里，在波士顿与洛杉矶之间以空中距离计算的最短通路是什么？在波士顿与洛杉矶之间什么样的航班组合的总飞行时间(即在空中的总时间，不包括航班之间的时间)最短？在这两个城市之间的最低费用是多少？在图 8-62 所示的计算机网络里，连接

旧金山的计算机与纽约的计算机所需要的最便宜的一组电话线是什么？哪一组电话线给出旧金山与纽约之间的通信的最短响应时间？哪一组电话线有最短的总距离？

与带权图有关的另外一个重要问题是：求访问完全图每个顶点恰好一次的、总长度最短的回路。这就是著名的旅行商问题，它求一位推销员应当以什么样的顺序来访问其路程上的每个城市恰好一次，使得他旅行的总距离最短。本节后面将讨论旅行商问题。

8.6.2 最短通路算法

求带权图里两个顶点之间的最短通路有几个不同的算法。下面将给出荷兰数学家 E·迪克斯特拉[⊖]在 1959 年所发现的一个解决无向带权图中最短通路问题的算法，其中所有的权都是正数。可以很容易地将它修改后来解决有向图里的最短通路问题。

在给出这个算法的形式化表示之前将给出一个启发性的例子。

例 1 在图 8-63 所示的带权图里，a 和 z 之间的最短通路的长度是多少？

解 虽然通过观察就容易求出最短通路，但是需要想出一些有助于理解迪克斯特拉算法的办法。要解决的问题就是：求从 a 到各个相继的顶点的最短通路，直到到达 z 为止。

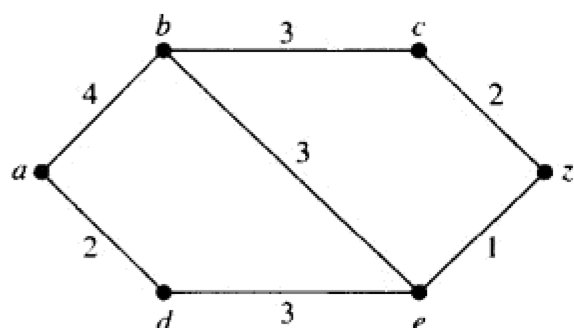


图 8-63 一个带权的简单图

从 a 开始、(直到到达终点为止)不包含除 a 以外的顶点的唯一通路是 a，b 和 a，d。因为 a，b 和 a，d 的长度分别为 4 和 2，所以 d 是与 a 最靠近的顶点。

可以通过查看(直到到达终点为止)只经过 a 和 d 的所有通路，来求出下一个最靠近 a 的顶点。到 b 的最短通路仍然是 a，b，长度为 4，而到 e 的最短通路是 a，d，e，长度为 5。所以，下一个与 a 最靠近的顶点是 b。

为了找出第三个与 a 最靠近的顶点，只需要检查(直到到达终点为止)只经过了 a，d 和 b 的那些通路。存在长度为 7 到 c 的通路，即 a，b，c，以及长度为 6 到 z 的通路，即 a，d，e，z。所以，z 是下一个与 a 最靠近的顶点，而且到 z 的最短通路的长度为 6。■

例 1 说明了在迪克斯特拉算法里使用的一般原理。注意通过检查就可能求出从 a 到 z 的最短通路。不过，无论对人还是对计算机来说，检查边数很多的图都是不切实际的。

现在将考虑一般问题：在无向连通简单带权图里，求出 a 与 z 之间的最短通路的长度。迪克斯特拉算法如下进行：求出从 a 到第一个顶点的最短通路的长度，从 a 到第二个顶点的最短通路的长度，依此类推，直到求出从 a 到 z 的最短通路的长度为止。

这个算法依赖于一系列的迭代。通过在每次迭代里添加一个顶点来构造出特殊顶点的集合。在每次迭代里完成一个标记过程。在这个标记过程里，用只包含特殊顶点的从 a 到 w 的最短通路的长度来标记 w。添加到特殊顶点集里的顶点是尚在集合之外的那些顶点中带有最小标记的顶点。

⊖ **爱德思葛·韦伯·迪克斯特拉(Edsger Wybe Dijkstra，1930—2002)** 迪克斯特拉出生在荷兰，他 20 世纪 50 年代早期在雷登大学学习理论物理时，就开始设计计算机程序。在 1952 年他意识到自己对程序设计比对物理学更感兴趣，于是他迅速地结束了对物理的学习，转而开始了程序员生涯，即使当时程序设计还没有被认为是一种职业。(在 1957 年，阿姆斯特丹当局拒绝接受他在结婚证上以“程序设计”作为职业。他只好把该项改成“理论物理学家”。)

迪克斯特拉一直是把程序设计作为一个科学学科的最有力的倡导者之一。他在下述领域做出了奠基性的贡献：操作系统，其中包括死锁避免；程序设计语言，其中包括结构化程序设计的概念；以及算法。在 1972 年迪克斯特拉接受了计算机学会的图灵奖，这是计算机科学里最有影响力的奖项之一。迪克斯特拉在 1973 年成为伯劳福斯研究员，又在 1984 年被任命为得克萨斯大学奥斯丁分校的计算机科学教授。

现在给出迪克斯特拉算法的细节。它首先用 0 标记 a 而用∞标记其余的顶点。用记号 $L_0(a)=0$ 和 $L_0(v)=\infty$ 表示在没有发生任何迭代之前的这些标记(下标 0 表示“第 0 次”迭代)。这些标记是从 a 到这些顶点的最短通路的长度，其中这些通路只包含顶点 a。(因为不存在从 a 到其他顶点的这种通路，所以∞是 a 与这样的顶点之间的最短通路的长度。)

迪克斯特拉算法是通过形成特殊顶点的集合来进行的。设 S_k 表示在标记过程 k 次迭代之后的特殊顶点集。首先让 $S_0=\varnothing$。集合 S_k 是通过把不属于 S_{k-1} 的带最小标记的顶点 u 添加到 S_{k-1} 里形成的。一旦把 u 添加到 S_k 里，就更新所有不属于 S_k 的顶点的标记，使得顶点 v 在第 k 个阶段的标记 $L_k(v)$ 是只包含 S_k 里顶点(即已有的特殊顶点再加上 u)的、从 a 到 v 的最短通路的长度。

设 v 是不属于 S_k 的一个顶点。为了更新 v 的标记，注意 $L_k(v)$ 是只包含 S_k 里顶点的从 a 到 v 的最短通路的长度。当利用下面的观察结果时，就可以有效地完成这个更新：只包含 S_k 里顶点的从 a 到 v 的最短通路，要么是只包含 S_{k-1} 里顶点(即不包括 u 在内的特殊顶点)的从 a 到 v 的最短通路，要么是在第 $k-1$ 阶段加上边 (u, v) 的从 a 到 u 的最短通路。换句话说，

$$L_k(a,v)=\min\{L_{k-1}(a,v), L_{k-1}(a,u)+w(u,v)\}$$

这个过程这样迭代：相继添加顶点到特殊顶点集里，直到添加 z 为止。当把 z 添加到特殊顶点集里时，它的标记就是从 a 到 z 的最短通路的长度。在算法 1 里给出迪克斯特拉算法。随后将证明这个算法的正确性。

算法 1 迪克斯特拉算法

```
procedure Dijkstra(G: 所有权都为正数的带权连通简单图)
{G 带有顶点 a=v0, v1, …, vn=z 和权 w(vi, vj)，其中若{vi, vj}不是 G 里的边，则 w(vi, vj)=∞}
for i := 1 to n
    L(vi) := ∞
L(a) := 0
S := ∅
{现在初始化标记，使得 a 的标记为 0 而所有其余标记为∞，S 是空集合}
while z ∉ S
begin
    u := 不属于 S 的 L(u) 最小的一个顶点
    S := S ∪ {u}
    for 所有不属于 S 的顶点 v
        if L(u)+w(u, v)<L(v) then L(v) := L(u)+w(u, v)
    {这样就给 S 里添加带最小标记的顶点，并且更新不在 S 里的顶点的标记}
end{L(z)=从 a 到 z 的最短通路的长度}
```

下面的例子说明迪克斯特拉算法是如何工作的。随后我们将证明这个算法总是产生带权图里两个顶点之间最短通路的长度。

例 2 用迪克斯特拉算法求图 8-64a 所示的带权图里顶点 a 与 z 之间最短通路的长度。

解 图 8-64 里显示了迪克斯特拉算法求 a 与 z 之间最短通路所用的步骤。在算法的每次迭代里，用圆圈圈起集合 S_k 里的顶点。对每次迭代都标明了只包含 S_k 里顶点的从 a 到每个顶点的最短通路。当圆圈圈到 z 时，算法终止。找到从 a 到 z 的最短通路是 a, c, b, d, e, z，长度为 13。 ■

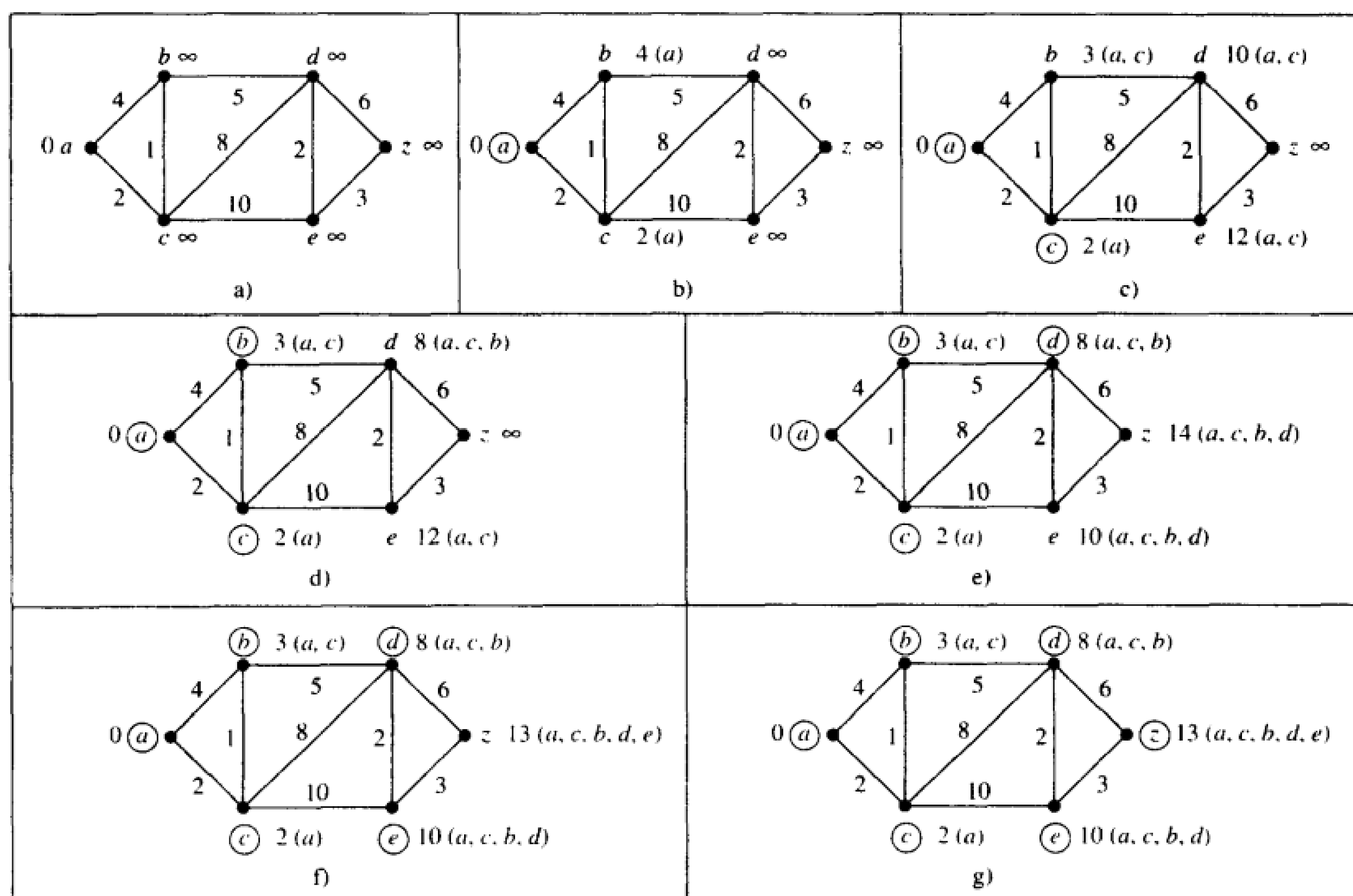


图 8-64 用迪克斯特拉算法求从 a 到 z 的最短通路

注意 在执行迪克斯特拉算法时，为了更便于在每步里跟踪顶点的标记，有时可以用一个表来代替，而不再对每步都重新画出这个图。

下一步，用归纳论证来证明迪克斯特拉算法产生无向连通带权图里两个顶点 a 与 z 之间最短通路的长度。用下列断言作为归纳假设：在第 k 次迭代里

(i) S 里的顶点 $v(v\neq 0)$ 的标记是从 a 到这个顶点的最短通路的长度。

(ii)不在 S 里的顶点的标记是（这个顶点自身除外）只包含 S 里顶点的、从 a 到这个顶点的最短通路的长度。

当 $k=0$ 时，在没有执行任何迭代之前，$S=\{a\}$，所以从 a 到除 a 外的顶点的最短通路的长度是∞。设 v 是在第 $k+1$ 次迭代里添加到 S 里的顶点，使得 v 是在第 k 次迭代结束时带最小标记的不在 S 里的顶点（在该顶点不唯一的情形里，可以采用带最小标记的任意顶点）。

根据归纳假设，可以看出在第 $k+1$ 次迭代之前，S 里的顶点都用从 a 出发的最短通路的长度来标记。另外，必须用从 a 到 v 的最短通路的长度来标记 v。假如情况不是这样，那么在第 k 次迭代结束时，就可能存在包含不在 S 里的顶点的、长度小于 $L_k(v)$ 的通路（因为 $L_k(v)$ 是在第 k 次迭代之后、只包含 S 里顶点的、从 a 到 v 的最短通路的长度）。设 u 是在这样的通路里不属于 S 的第一个顶点。则存在一条只包含 S 里顶点的、从 a 到 u 的长度小于 $L_k(v)$ 的通路。这与对 v 的选择相矛盾。因此，在第 $k+1$ 次迭代结束时(i)成立。

设 u 是在第 $k+1$ 次迭代之后不属于 S 的一个顶点。只包含 S 里顶点的从 a 到 u 的最短通路要么包含 v、要么不包含 v。若它不包含 v，则根据归纳假设，它的长度是 $L_k(u)$。若它确实包含 v，则它必然是这样组成的：一条只包含 S 里除 v 之外的顶点的、从 a 到 v 的、具有最短可能长

度的通路，后面接着从 v 到 u 的边。在这种情形里它的长度是 $L_k(v)+w(v, u)$。这样就证明了(ii)为真，因为 $L_{k+1}(u)=\min\{L_k(u), L_k(v)+w(v, u)\}$。

已经证明了定理 1。

定理 1 迪克斯特拉算法求出连通简单无向带权图里两个顶点之间最短通路的长度。

现在可以估计迪克斯特拉算法的计算复杂性(就加法和比较而言)。这个算法使用的迭代次数不超过 $n-1$ 次，因为在每次迭代里添加一个顶点到特殊顶点集里。若可以估计每次迭代所使用的运算次数，则大功告成了。可以用不超过 $n-1$ 次比较来找出不在 S_k 里的带最小标记的顶点。于是我们使用一次加法和一次比较来更新不在 S_k 中的每一个顶点的标记，所以在每次迭代里的运算不超过 $2(n-1)$次，因为在每次迭代里要更新的标记不超过 $n-1$ 个。因为迭代次数不超过 $n-1$ 次，每次迭代的运算次数不超过 $2(n-1)$次，所以有下面的定理。

定理 2 迪克斯特拉算法使用 $O(n^2)$次运算(加法和比较)来求出连通简单无向带权图里两个顶点之间最短通路的长度。

8.6.3 旅行商问题

现在讨论与带权图有关的一个重要问题。考虑下面的问题：一位旅行商想要访问 n 个城市中每个城市恰好一次，并且返回到出发点。例如，假定这个旅行商想要访问底特律、托莱多、萨吉诺、格兰特·拉皮滋以及卡拉玛祖(见图 8-65)。他应当以什么顺序访问这些城市以便旅行总距离最短？为了解决这个问题，可以假定旅行商从底特律出发(因为这个城市必须是回路的一部分)，并且检查他访问其余四个城市然后返回底特律的所有可能方式(从别处出发将产生相同的回路)。这样的回路总共有 24 条，但是因为往返路程距离相同，所以在求最短总距离时，只需要考虑 12 条不同的回路即可。列出这 12 条不同回路和每条回路旅行的最短总距离。从表 8-6 可以看出，回路底特律-托莱多-卡拉玛祖-格兰特·拉皮滋-萨吉诺，对应 458 英里的最短总距离。

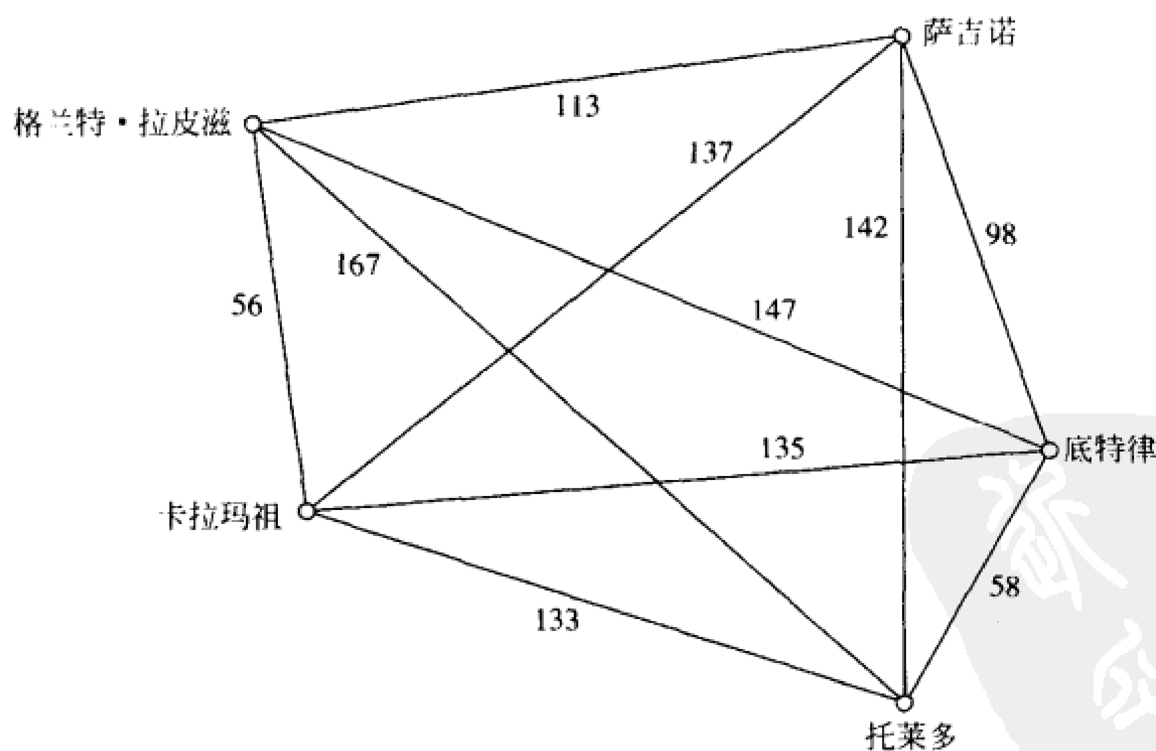


图 8-65 说明五个城市之间距离的图

上面描述了旅行商问题的一个实例。旅行商问题求带权完全无向图里访问每个顶点恰好一次并且返回出发点的总权数最小的回路。这等价于求完全图里总权数最小的哈密顿回路，因为在回路里访问每个顶点恰好一次。

表 8-6

路　线	总距离(英里)
底特律—托莱多—格兰特·拉皮滋—萨吉诺—卡拉玛祖—底特律	610
底特律—托莱多—格兰特·拉皮滋—卡拉玛祖—萨吉诺—底特律	516
底特律—托莱多—卡拉玛祖—萨吉诺—格兰特·拉皮滋—底特律	588
底特律—托莱多—卡拉玛祖—格兰特·拉皮滋—萨吉诺—底特律	458
底特律—托莱多—萨吉诺—卡拉玛祖—格兰特·拉皮滋—底特律	540
底特律—托莱多—萨吉诺—格兰特·拉皮滋—卡拉玛祖—底特律	504
底特律—萨吉诺—托莱多—格兰特·拉皮滋—卡拉玛祖—底特律	598
底特律—萨吉诺—托莱多—卡拉玛祖—格兰特·拉皮滋—底特律	576
底特律—萨吉诺—卡拉玛祖—托莱多—格兰特·拉皮滋—底特律	682
底特律—萨吉诺—格兰特·拉皮滋—托莱多—卡拉玛祖—底特律	646
底特律—格兰特·拉皮滋—萨吉诺—托莱多—卡拉玛祖—底特律	670
底特律—格兰特·拉皮滋—托莱多—萨吉诺—卡拉玛祖—底特律	728

最直截了当的求解旅行商问题实例的方式是检查所有可能的哈密顿回路，并且挑出总权数最小的一条回路。若在图中有 n 个城市，则为了求解这个问题，得检查多少条回路？一旦选定了出发点，需要检查的不同的哈密顿回路就有$(n-1)!$ 条，因为第二个顶点有 $n-1$ 种选择，第三个顶点有 $n-2$ 种选择，依此类推。因为可以用相反顺序来经过一条哈密顿回路，所以只需要检查$(n-1)!/2$ 条回路来求出答案。注意$(n-1)!/2$ 增长得极快。当只有几十个顶点时，试图用这种方式来解决旅行商问题就是不切实际的。例如，假如有 25 个顶点，那么就不得不考虑总共 $24!/2$(约为 3.1×10^{23})条不同的哈密顿回路。假定检查每条哈密顿回路只花费 1 纳秒(10^{-9}秒)，那么就需要大约一千万年才能求出这个图里长度最短的一条哈密顿回路。

因为旅行商问题在实践和理论上都具有重要意义，所以已经投入了巨大的努力来设计解决它的有效算法。不过，还没有已知的解决这个问题的多项式最坏情形时间复杂性的算法。进一步地，假如这种算法找到了，那么许多其他难题(比如在第 1 章里讨论过的确定 n 个变元的命题公式是否重言式)也可以迎刃而解。这个结果是从 NP 完全性理论得出的(关于这个理论的更多信息请参考[GaJo79])。

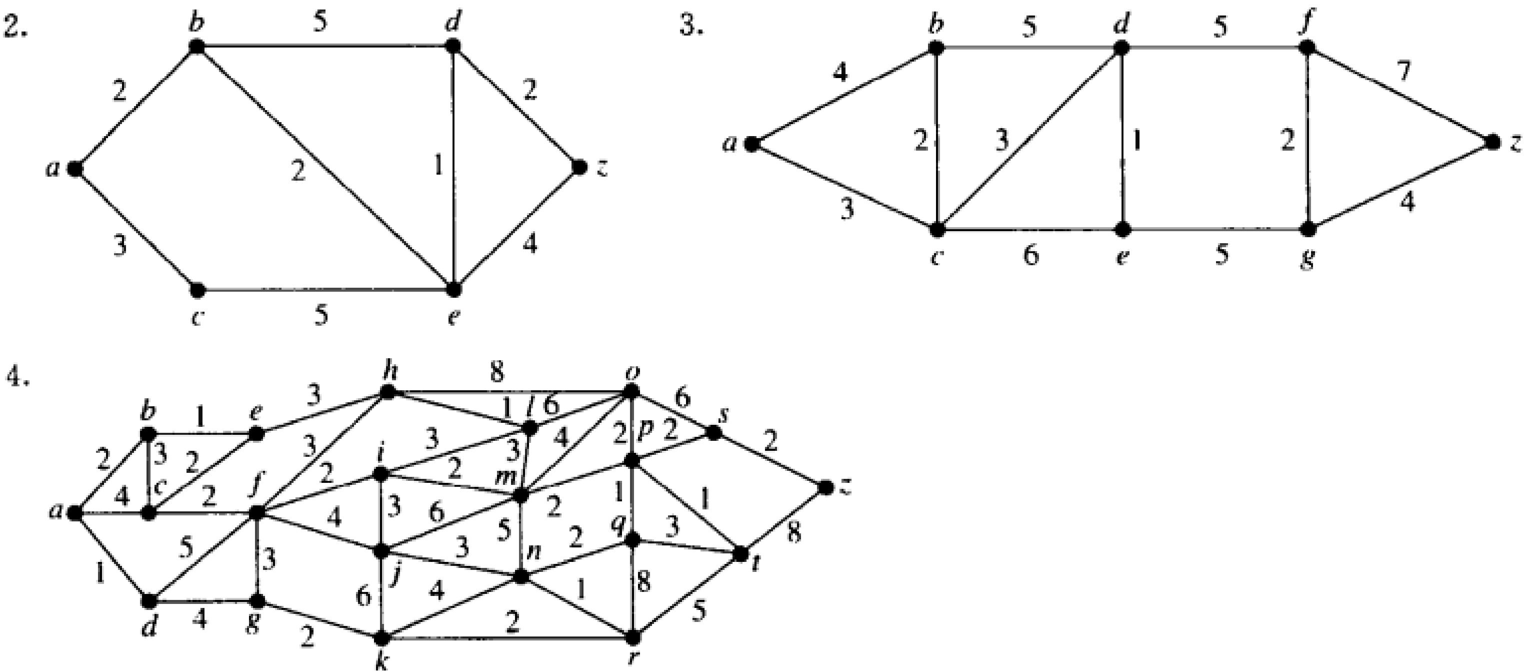
当有许多需要访问的顶点时，解决旅行商问题的实际方法是使用*近似算法*。近似算法是这样的算法，它们不必产生问题的精确解，取而代之的是保证产生接近精确解的解。即它们可能产生带总权数 W' 的哈密顿回路，使得 $W\leqslant W'\leqslant cW$，其中 W 是精确解的总长度，而 c 是一个常数。例如，存在多项式最坏情形时间复杂性算法使得 $c=3/2$。对于一般带权图和每个正实数 k 来说，总是产生至多 k 倍于最优解的解的算法还是未知。假如这样的算法存在，那就可能证明 P 类等于 NP 类，这是关于算法复杂性的最著名的未解决问题(参见 2.3 节)。

在实际中，已经研究出这样的算法，它们可以只用几分钟的机时，就解决多达 1000 个顶点的旅行商问题，误差在精确解的 2%之内。关于旅行商问题的更多信息，包括历史、应用和算法等，见《*Applications of Discrete Mathematics*》(离散数学的应用)[MiRo91]中关于这个主题的那一章。

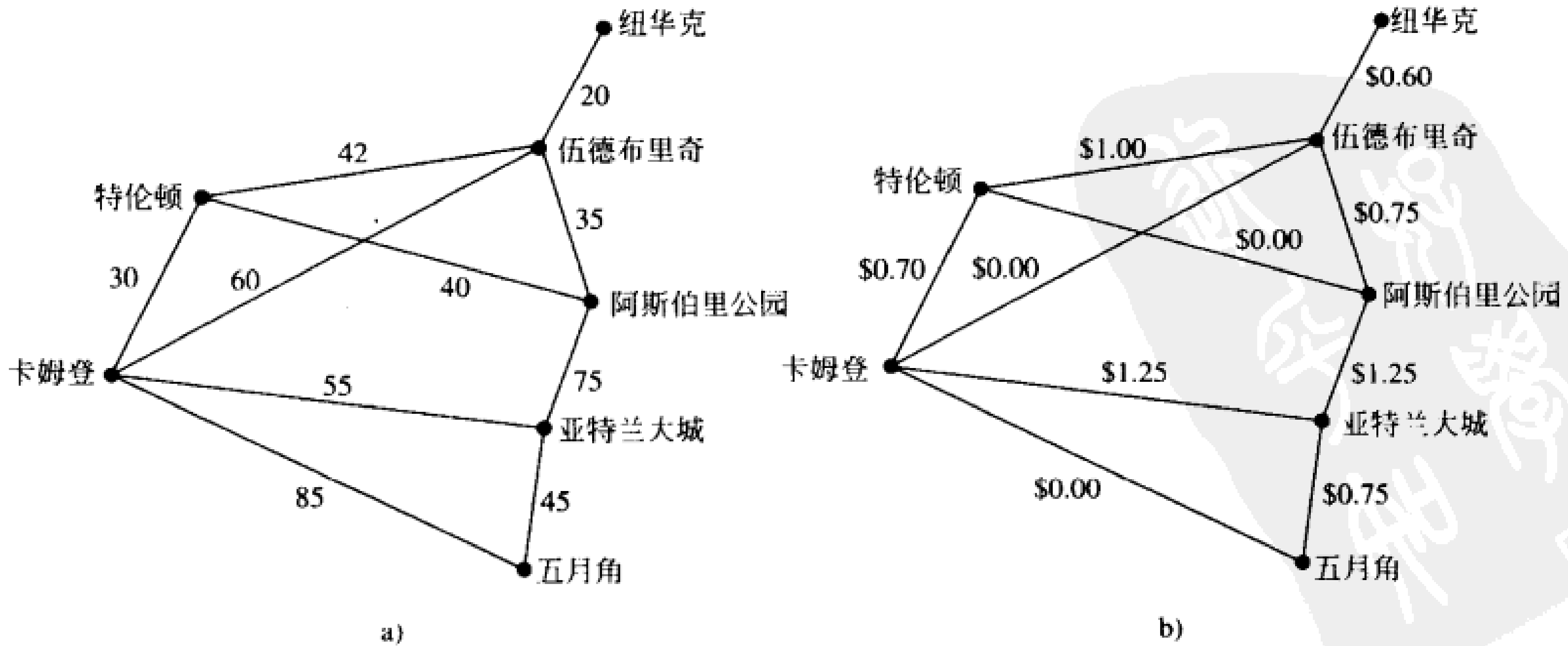
练习

1. 对下列关于地铁系统的每个问题，描述一个可以用来解决这个问题的带权图模型。
   a)在两站之间旅行所需要的最短时间是什么？
   b)从一站到达另外一站所经过的最短距离是什么？
   c)若把各站之间的票价求和就得出总票价，则在两站之间的最低票价是什么？

在练习 2～4 里，求给定带权图里在 a 与 z 之间的最短通路的长度。



5. 在练习 2～4 里的每个带权图中，在 a 与 z 之间的最短通路是什么？
6. 在练习 3 的带权图里，求下列成对顶点之间的最短通路的长度。
   a) a 和 d　　b) a 和 f　　c) c 和 f　　d) b 和 z
7. 在练习 3 的带权图里，求练习 6 里的成对顶点之间的最短通路。
8. 在图 8-61 所示的航线系统里，求下列每对城市之间的最短通路(以英里表示)。
   a) 纽约与洛杉矶　　b) 波士顿与旧金山　　c) 迈阿密与丹佛　　d) 迈阿密与洛杉矶
9. 利用图 8-61 所示的飞行时间，求连接练习 8 中成对城市的总飞行时间最短的航班组合。
10. 利用图 8-61 所示的票价，求连接练习 8 中成对城市的票价最低的航班组合。
11. 在图 8-62 所示的通信网络里，求下列每对城市的计算机中心之间的(距离)最短路线。
    a) 波士顿与洛杉矶　　b) 纽约与旧金山　　c) 达拉斯与旧金山　　d) 丹佛与纽约
12. 利用在图 8-62 里给出的响应时间，求在练习 11 中成对的计算机中心之间响应时间最短的路线。
13. 利用在图 8-62 里给出的租费，求在练习 11 中成对的计算机中心之间月租费最便宜的路线。
14. 解释把无向图里两个顶点之间边数最少的通路当作带权图里的最短通路来求解的过程。
15. 推广求带权简单有向连通图里两个顶点之间最短通路的迪克斯特拉算法，以便求出在顶点 a 与图中其余每个顶点之间的最短通路的长度。
16. 推广求带权简单有向连通图里两个顶点之间最短通路的迪克斯特拉算法，以便构造出在这些顶点之间的最短通路。
17. 在下图里的带权图说明新泽西的一些主要道路。图 a 说明这些道路上的城市之间的距离；图 b 说明通行费。



a)利用这些道路，求在纽华克与卡姆登之间、以及在纽华克与五月角之间的距离最短的路线。

b)利用给出的道路图，求在本题部分(a)里成对城市之间就总通行费而言最便宜的路线。

18. 若各边的权都是不同的，则在带权图里两个顶点之间的最短通路是否唯一？

19. 哪些应用必须求出带权图里两个顶点之间的最长简单通路的长度？

20. 什么是图 8-64 的带权图里在 a 与 z 之间的最长简单通路的长度？在 c 与 z 之间呢？

可以用*弗洛伊德算法*来求出带权连通简单图里所有顶点对之间最短通路的长度。不过，不能用这个算法来构造出最短通路(在下面，把无穷权赋给任何一对不被图中的边所连接的顶点上)。

算法 2 弗洛伊德算法

procedure $Floyed(G$：带权简单图)
{G 有顶点 v_1，v_2，…，v_n 和权 $w(v_i, v_j)$，其中若(v_i，v_j)不是边，则 $w(v_i, v_j)=\infty$}
for $i:=1$ **to** n
    for $j:=1$ **to** n
        $d(v_i, v_j):=w(v_i, v_j)$
for $i:=1$ **to** n
    for $j:=1$ **to** n
        for $k:=1$ **to** n
            if $d(v_j, v_i)+d(v_i, v_k)<d(v_j, v_k)$
                then $d(v_j, v_k):=d(v_j, v_i)+d(v_i, v_k)$
{$d(v_i, v_j)$是在 v_i 与 v_j 之间的最短通路的长度}

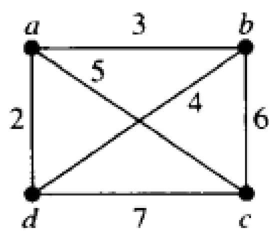
21. 用弗洛伊德算法求图 8-64a 中带权图里所有顶点对之间的距离。

*22. 证明：弗洛伊德算法确定出带权简单图里所有顶点对之间的最短距离。

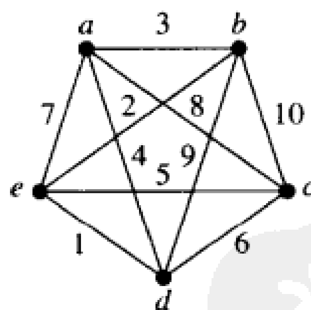
*23. 给出弗洛伊德算法为了确定在带有 n 个顶点的带权简单图里、所有顶点对之间的最短距离而使用的运算(比较和加法)次数的大 O 估计。

*24. 证明：若边有负权，则迪克斯特拉算法或许不能给出正确答案。

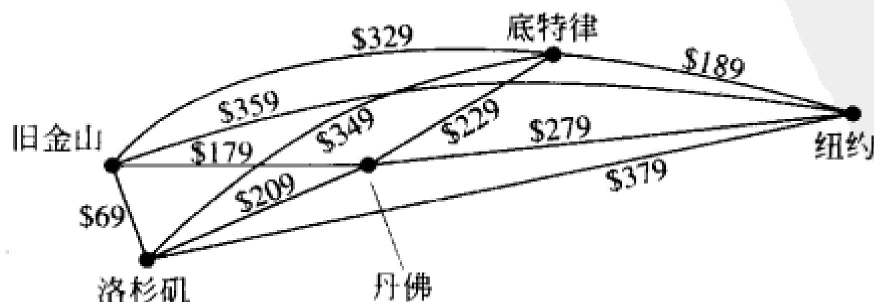
25. 通过求出所有哈密顿回路的总权数并且确定出总权数最小的回路，来解决下图的旅行商问题。



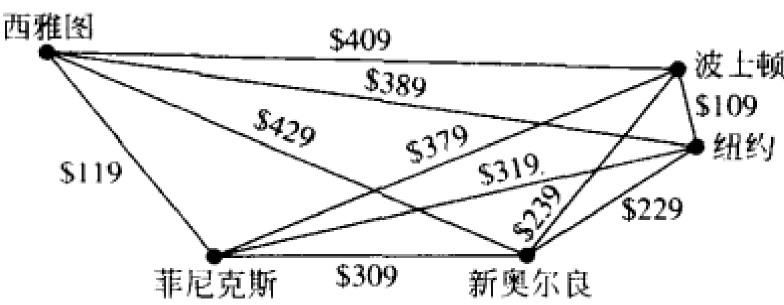
26. 通过求出所有哈密顿回路的总权数并且确定出总权数最小的回路，来解决下图的旅行商问题。



27. 求访问下图里每个城市的机票总价最低的路线，其中边上的权是在这两个城市之间的航班所提供的最低票价。



28. 求访问下图里每个城市的机票总价最低的路线，其中边上的权是在这两个城市之间的航班所提供的最低票价。



29. 构造一个带权无向图，使得对于访问某些顶点超过一次的回路来说，访问每个顶点至少一次的回路的总权数是最小的（提示：存在带三个顶点的例子）。
30. 证明：求访问带权图每个顶点至少一次的总权数最小的回路的问题，可以归约为求访问带权图每个顶点恰好一次的总权数最小的回路的问题。这样做的方法是：构造一个新的带权图，它与原图有相同的顶点和边，但是连接顶点 u 和 v 的边的权却等于在原图里从 u 到 v 的通路的最小总权数。

8.7 可平面图

8.7.1 引言

考虑把三座房屋与三种设施每种都连接起来的问题，如图 8-66 所示。是否有可能这样来连接这些房屋与设施，使得在这样的连接里不发生交叉？这个问题可以用完全偶图 $K_{3,3}$ 来建模。原来的问题可以重新叙述为：能否在平面里画出 $K_{3,3}$，使得没有两条边发生交叉？

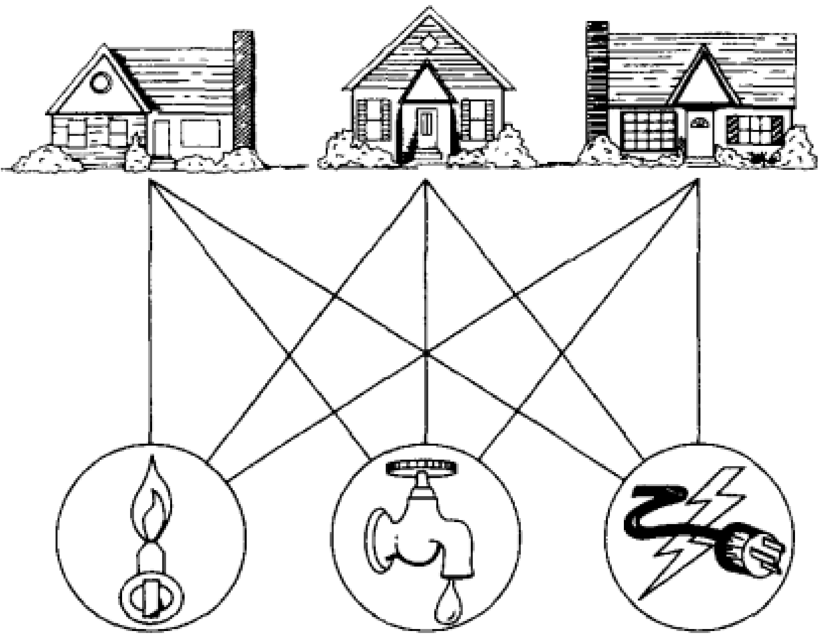


图 8-66 三座房屋与三种设施

本节将研究能否在平面里让边不交叉地画出一个图的问题。特别是，将回答这个房屋与设施的问题。

图的表示方式有许多种。何时有可能至少求出一种方式来在平面里表示这个图而让边没有任何交叉？

定义 1 若可以在平面里画出一个图而边没有任何交叉（其中边的交叉是表示边的直线或弧线在它们的公共端点以外的地方相交），则这个图是*可平面的*。这种画法称为这个图的*平面表示*。

即使通常带交叉地画出了一个图，这个图也仍然可能是可平面的，这是因为有可能以不同的方式来不带交叉地画出这个图。

例 1 K_4（如图 8-67 所示，有两条边交叉）是可平面吗？

解 K_4 是可平面的，因为可以不带交叉地画出它，如图 8-68 所示。 ■

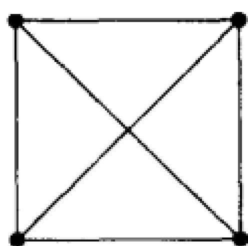


图 8-67 图 K_4

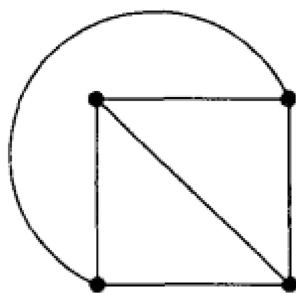


图 8-68 不带交叉的图 K_4

例 2 图 8-69 所示的 Q_3 是可平面吗？

解 Q_3是可平面的，因为可以画出它而没有任何边交叉，如图 8-70 所示。 ■

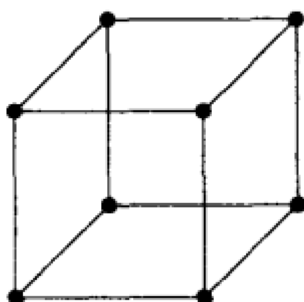


图 8-69 图 Q_3

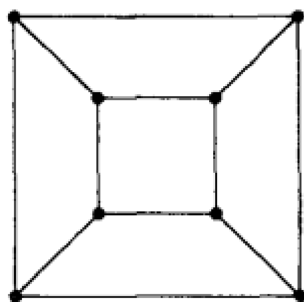


图 8-70 Q_3 的一种平面表示

可以通过显示一种平面表示来证明一个图是可平面的。更难的是证明一个图是非可平面的。下面给出一个例子说明如何以一种特别的方法来做到这一点。

例 3 图 8-71 所示的 $K_{3,3}$是可平面吗？

解 在平面里画出 $K_{3,3}$而没有边交叉，任何这样的尝试都注定是失败的。现在说明这是为什么。在 $K_{3,3}$的任何平面表示里，顶点 v_1 和 v_2 都必须同时与 v_4 和 v_5 连接。这四条边所形成的封闭曲线把平面分割成两个区域 R_1 和 R_2，如图 8-72a 所示。顶点 v_3 属于 R_1 或 R_2。当 v_3 属于闭曲线的内部 R_2 时，在 v_3 和 v_4 之间以及在 v_3 和 v_5 之间的边，把 R_2 分割成两个区域 R_{21} 和 R_{22}，如图 8-72b 所示。

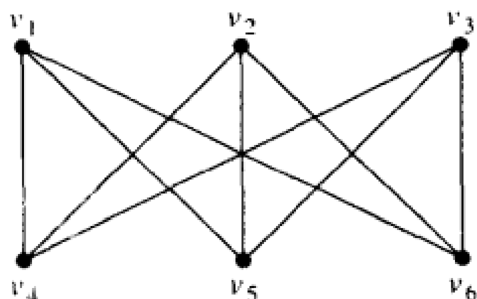


图 8-71 图 $K_{3,3}$

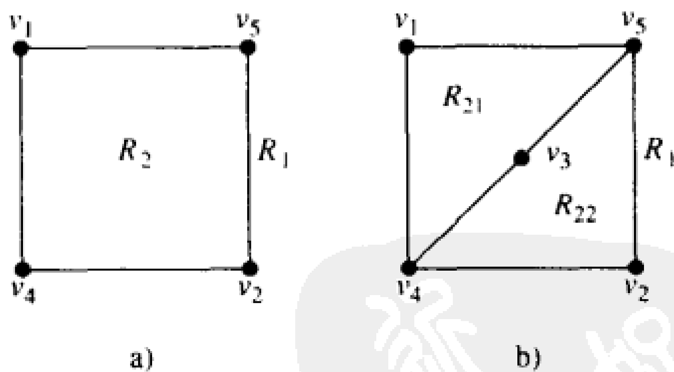


图 8-72 证明 $K_{3,3}$是非可平面的

下一步，已经没有办法来不交叉地放置最后一个顶点 v_6。因为若 v_6 属于 R_1，则不能不带交叉地画出 v_6 和 v_3 之间的边。若 v_6 属于 R_{21}，则不能不带交叉地画出 v_2 和 v_6 之间的边。若 v_6 属于 R_{22}，则不能不带交叉地画出 v_1 和 v_6之间的边。

当 v_3 属于 R_1 时，可以使用类似的论证。请读者来完成这个论证（见本节末的练习 10）。所以 $K_{3,3}$是非可平面的。 ■

例 3 解决了在本节开头所描述的设施与房屋的问题。不能在平面里连接这三座房屋与三种设

施而不发生交叉。可以用类似的论证来证明 K_5 是非可平面的(见本节末的练习 11)。

图的平面性在电子电路的设计中有重要作用。可以用图来为电路建立模型，用顶点表示电路的器件并且用边表示器件之间的连接。如果表示一个电路的图是可平面的，则可以把这个电路无连接交叉地印刷在单个电路板上。当这个图不是可平面的时，就必须求助于成本更高的选项。例如，可以把表示电路的图的顶点划分成可平面子图。然后使用多层来构造这个电路(参见练习 30 前面的说明来了解图的厚度)。每当连接交叉时就可以用绝缘线来构造电路。在这种情况下，重要的是以尽可能少的交叉来画出这个图(参见练习 26 前面的说明来了解图的交叉数)。

8.7.2 欧拉公式

一个图的可平面表示把平面分割成一些面，包括一个无界的面。例如，图 8-73 所示的图的可平面表示把平面分割成 6 个面并加以标记。欧拉证明过一个图的所有可平面表示都把平面分割成相同数目的面。他是这样证明的：求出可平面图的面数、顶点数以及边数之间的关系。

定理 1 欧拉公式。设 G 是带 e 条边和 v 个顶点的连通可平面简单图。设 r 是 G 的可平面表示里的面数。则 $r=e-v+2$。

证 首先规定 G 的可平面表示。将要这样证明定理：构造一系列子图 G_1，G_2，…，$G_e=G$，相继地在每个阶段添加一条边。用下面的归纳定义来这样做。任意地选择一条 G 的边来获得 G_1。这样通过 G_{n-1} 获得 G_n：任意地添加一条与 G_{n-1} 里顶点相关联的边，若与这条边关联的另一个顶点还不在 G_{n-1} 里，则添加这个顶点。这样的构造是可能的，因为 G 是连通的。在添加 e 条边之后就获得 G。设 r_n，e_n 和 v_n 分别表示 G 的可平面表示所引导出的 G_n 的可平面表示的面数、边数和顶点数。

现在通过归纳来进行证明。对 G_1 来说，关系 $r_1=e_1-v_1+2$ 为真，因为 $e_1=1$，$v_1=2$，而 $r_1=1$。这种情形如图 8-74 所示。

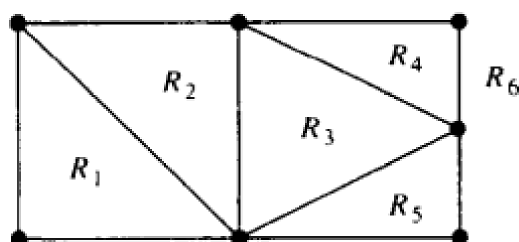


图 8-73 图的可平面表示的面

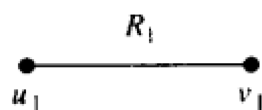


图 8-74 欧拉公式的证明中的基础情形

现在假定 $r_n=e_n-v_n+2$。设 $\{a_{n+1}, b_{n+1}\}$ 是为了获得 G_{n+1} 而添加到 G_n 上的边。有两种情形需要考虑。在第一种情形里，a_{n+1} 和 b_{n+1} 都已经在 G_n 里了。这两个顶点必然是在一个公共面 R 的边界上，否则就不可能把边 $\{a_{n+1}, b_{n+1}\}$ 添加到 G_n 里而没有两条边交叉(并且 G_{n+1} 是可平面的)。这条新边的添加把 R 分割成两个面。所以，在这种情形里，$r_{n+1}=r_n+1$，$e_{n+1}=e_n+1$，而且 $v_{n+1}=v_n$。因此，关系到面数、边数、顶点数的公式两边都恰好增加 1，所以这个公式仍然为真。换句话说，$r_{n+1}=e_{n+1}-v_{n+1}+2$。在图 8-75a 里说明这种情形。

在第二种情形里，新边的两个顶点之一还不在 G_n 里。假定 a_{n+1} 在 G_n 里但是 b_{n+1} 不在 G_n 里。添加这条新边不产生任何新的面，因为 b_{n+1} 必然是在边界上有 a_{n+1} 的一个面里。所以，$r_{n+1}=r_n$。另外，$e_{n+1}=e_n+1$ 且 $v_{n+1}=v_n+1$。关系到面数、边数、顶点数的公式两边都保持相等，所以这个公式仍然为真。换句话说，$r_{n+1}=e_{n+1}-v_{n+1}+2$。在图 8-75b 里说明这种情形。

已经完成了归纳论证。因此对所有 n 来说，都有 $r_n=e_n-v_n+2$。因为原图是在添加了 e 条

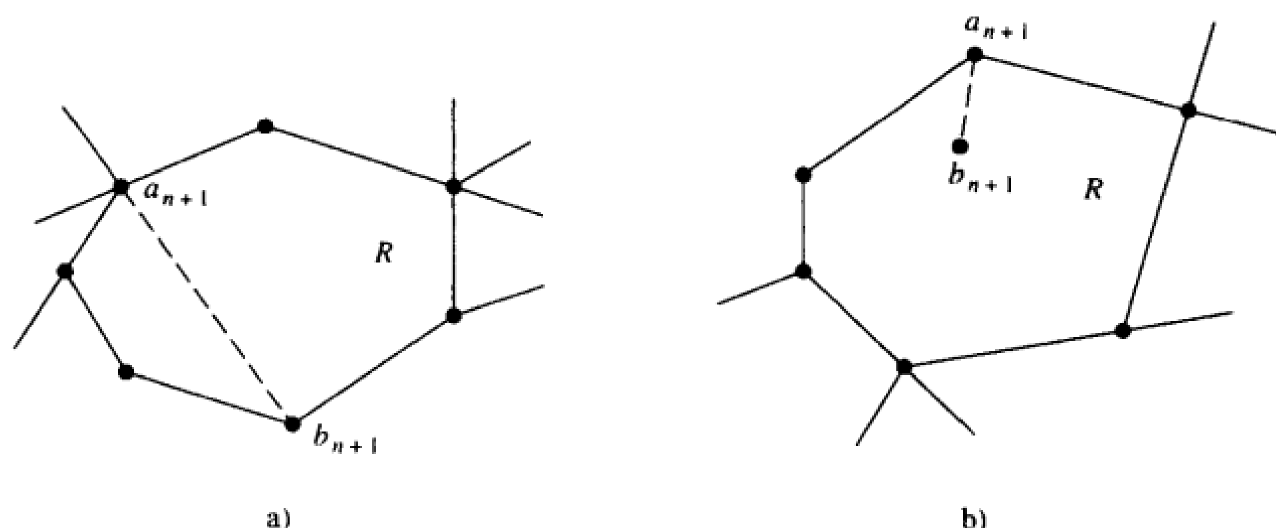


图 8-75 添加一条边到 G_n 来产生 G_{n+1}

边之后所获得的图 G_e，所以这个定理为真。 □

例 4 解释欧拉公式。

例 4 假定连通可平面简单图有 20 个顶点，每个顶点度都为 3。这个可平面图的平面表示把平面分割成多少个面?

解 这个图有 20 个顶点，每个顶点的度都为 3，所以 $v=20$。因为这些顶点的度之和 $3v=3\cdot 20=60$ 等于边数的两倍 $2e$，所以有 $2e=60$ 或 $e=30$。所以，根据欧拉公式，面数是

$$r=e-v+2=30-20+2=12$$ ■

可以用欧拉公式来建立可平面图所必须满足的一些不等式。在下面的推论里给出一个这样的不等式。

推论 1 若 G 是带 e 条边和 v 个顶点的连通可平面简单图，其中 $v\geqslant 3$，则 $e\leqslant 3v-6$。

在证明推论 1 之前先用它证明下面这个有用的结论。

推论 2 若 G 是连通可平面简单图，则 G 有度数不超过 5 的顶点。

证 如果 G 有 1 个或 2 个顶点，则结果为真。如果 G 至少有 3 个顶点，则根据推论 1 知道 $e\leqslant 3v-6$，所以 $2e\leqslant 6v-12$。假如每个顶点的度数至少是 6，则由于 $2e=\sum_{v\in V}\deg(v)$（根据握手定理），所以就有 $2e\geqslant 6v$。但是这与 $2e\leqslant 6v-12$ 相矛盾。所以必定存在度数不超过 5 的顶点。□

推论 1 的证明是基于面的次数的概念，它定义为这个面的边界上的边数。当一条边在边界上出现两次(所以当描画边界时就描画它两次)时，它贡献的次数是 2。图 8-76 里显示出面的次数。

现在可以给出推论 1 的证明了。

证 画在平面里的连通可平面简单图把平面分割成面，比如说 r 个面。每个面的次数至少为 3。(因为这里所讨论的图都是简单图，不允许带有可能产生 2 次面的多重边，或者可能产生 1 次面的环。)特别地，注意无界的面的次数至少为 3，因为在图里至少有三个顶点。

注意各面的次数之和恰好是图中边数的两倍，因为每条边都在面的边界上出现两次(或者在两个不同面里，或者两次在相同面里)。因为每个面都有大于或等于 3 的次数，所以

$$2e=\sum_{\text{所有区域}R}\deg(R)\geqslant 3r$$

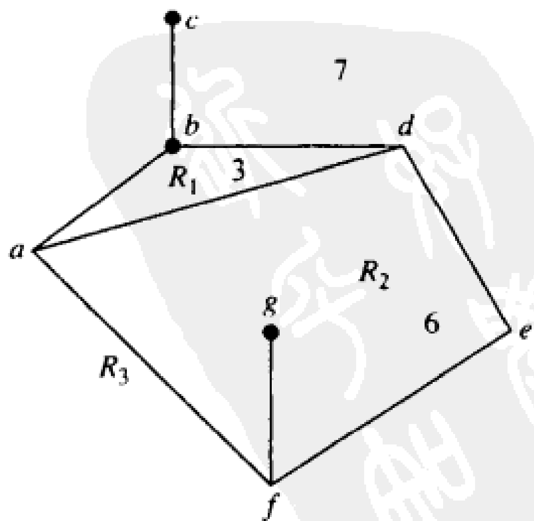


图 8-76 面的次数

因此，

$$(2/3)e \geqslant r$$

利用 $r=e-v+2$(欧拉公式)，就得到

$$e-v+2 \leqslant (2/3)e$$

所以 $e/3 \leqslant v-2$。这样就证明了 $e \leqslant 3v-6$。 □

可以用这个推论来证明 K_5 是非可平面的。

例 5 用推论 1 证明：K_5 是非可平面的。

解 图 K_5 有 5 个顶点和 10 条边。不过，对这个图来说，不满足不等式 $e \leqslant 3v-6$，因为 $e=10$ 和 $3v-6=9$。因此，K_5 不是可平面的。 ■

前面已经证明了 $K_{3,3}$ 不是可平面的。不过，注意这个图有 6 个顶点和 9 条边。这意味着满足不等式 $e=9 \leqslant 12=3 \cdot 6-6$。所以，满足不等式 $e \leqslant 3v-6$ 的事实并不意味着一个图是可平面的。不过，可以利用下面定理 1 的推论来证明 $K_{3,3}$ 不是可平面的。

推论 3 若连通可平面简单图有 e 条边和 v 个顶点，$v \geqslant 3$ 并且没有长度为 3 的回路，则 $e \leqslant 2v-4$。

推论 3 的证明类似于推论 1 的证明，不同之处在于，在这种情形里，没有长度为 3 的回路的事实蕴含着面的次数必然至少为 4。把这个证明的细节留给读者(见本节末的练习 15)。

例 6 用推论 3 来证明 $K_{3,3}$ 是非可平面的。

解 因为 $K_{3,3}$ 没有长度为 3 的回路(容易看出这一点，因为它是偶图)，所以可以使用推论 3。$K_{3,3}$ 有 6 个顶点和 9 条边。因为 $e=9$ 和 $2v-4=8$，所以推论 3 证明 $K_{3,3}$ 是非可平面的。 ■

8.7.3 库拉图斯基定理

我们已经看到 $K_{3,3}$ 和 K_5 都不是可平面的。显然，若一个图包含这两个图作为子图，则它不是可平面的。另外，所有非可平面图必然包含一个子图，它是可以利用某些允许的操作从 $K_{3,3}$ 或 K_5 来获得的。

若一个图是可平面的，则通过删除一条边 $\{u, v\}$ 并且添加一个新顶点 w 和两条边 $\{u, w\}$ 与 $\{w, v\}$，所获得的任何图也是可平面的。这样的操作称为*初等细分*。若可以从相同的图通过一系列初等细分来获得图 $G_1=(V_1, E_1)$ 和图 $G_2=(V_2, E_2)$，则它们称为是*同胚的*。图 8-77 所示的三个图都是同胚的，因为它们都可以从第一个图通过初等细分来获得(读者应当确定出从 G_1 获得 G_2 和 G_3 所需要的初等细分的序列)。

波兰数学家库拉图斯基[⊖] 在 1930 年证明了下面的定理，这个定理用图同胚的概念刻画了可平面图。

⊖ **卡兹米尔兹·库拉图斯基(Kazimierz Kuratowski, 1896—1980)** 库拉图斯基是华沙一位著名律师的儿子。他在华沙上了高中。1913 年到 1914 年，他在苏格兰的格拉斯哥学习，但是，在第一次世界大战爆发后无法返回那里继续他的学业。1915 年他进入华沙大学，投身于波兰爱国学生运动。他在 1919 年发表了第一篇论文，并且在 1921 年获得博士学位。他是以华沙数学学派著称的小组里的活跃成员，以集合论基础和拓扑学为研究领域。他被勒沃工业大学聘为副教授，并在那里呆了 7 年，与重要的波兰数学家巴拿赫和乌拉姆合作。在 1930 年还在勒沃时，库拉图斯基完成了刻画可平面图的工作。

1934 年他作为教授，回到华沙大学。他一直从事研究和教学工作，直到第二次世界大战开始。在战争期间，为避免受到迫害，他起用假名来隐藏并且在秘密的华沙大学教书。战后他帮助恢复波兰的数学研究工作，担任波兰国家数学研究所主任。他写过 180 多篇论文，以及三本被广泛使用的教科书。

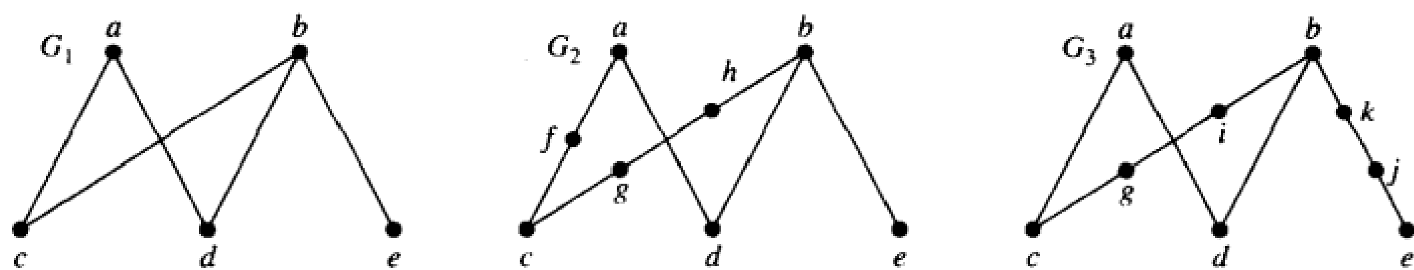


图 8-77 同胚图

定理 2 一个图是非可平面的当且仅当它包含一个同胚于 $K_{3,3}$或 K_5 的子图。

显然，一个包含着同胚于 $K_{3,3}$或 K_5 的子图的图是非可平面的。不过，相反方向的命题——即每个非可平面图都包含一个同胚于 $K_{3,3}$或 K_5 的子图——证明起来是复杂的，因而不在这里给出。下面的例子说明如何使用库拉图斯基定理。

例 7 确定图 8-78 所示的图 G 是否是可平面的。

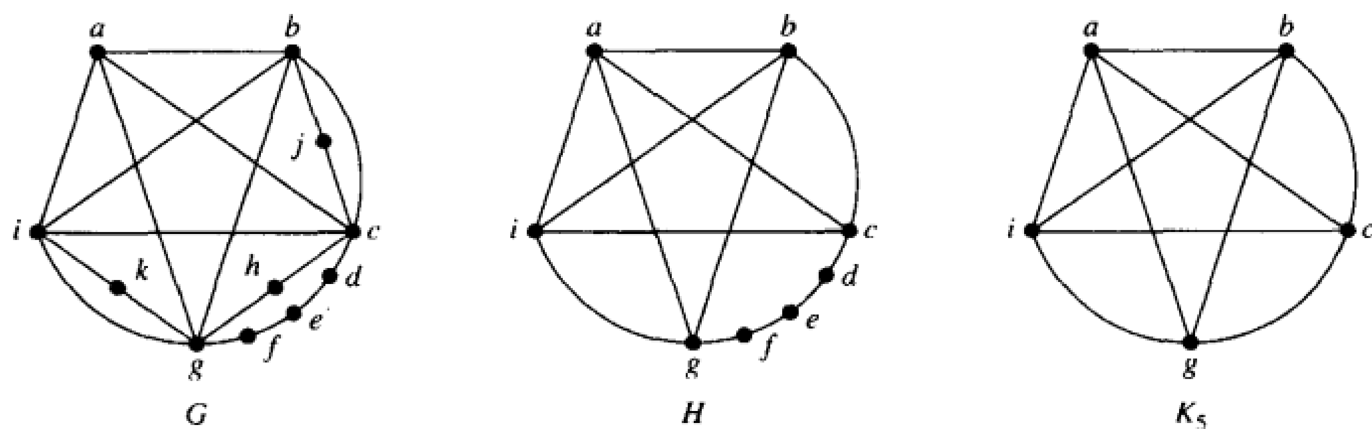


图 8-78 无向图 G、同胚于 K_5 的子图 H 以及 K_5

Extra Examples **解** G 有同胚于 K_5 的子图 H。H 是这样获得的：删除 h，j 和 k 以及所有与这些顶点关联的边。H 是同胚于 K_5 的，因为从 K_5(带有顶点 a，b，c，g 和 i)通过一系列初等细分，添加顶点 d，e 和 f 就可以获得 H(读者应当构造出这样的一系列初等细分)。因此，G 是非可平面的。■

例 8 在图 8-79a 中所示的彼得森图是可平面吗？(丹麦数学家朱利乌斯·彼得森在 1891 年介绍过这个图；它常用来说明关于图的各种性质的理论。)

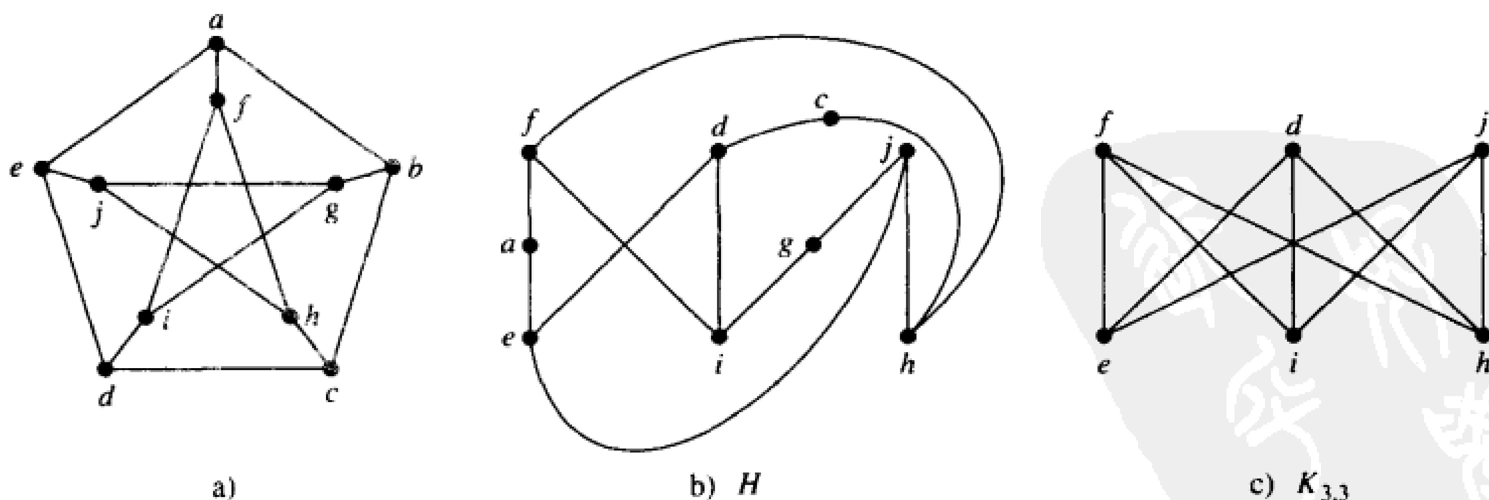


图 8-79 a)彼得森图，b)同胚于 $K_{3,3}$ 的子图 H，c)$K_{3,3}$

解 彼得森图的子图 H 是这样获得的：删除 b 和以 b 作为端点的三条边，如图 8-77b 所示，它是同胚于带有顶点集合$\{f, d, j\}$和$\{e, i, h\}$的 $K_{3,3}$的，这是因为可以通过一系列初等细分来

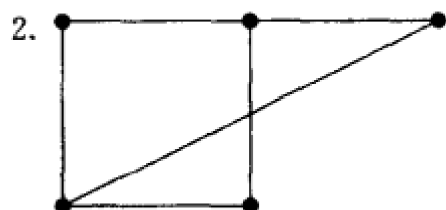
获得它，删除$\{d,h\}$并且添加$\{c,h\}$和$\{c,d\}$，删除$\{e,f\}$并且添加$\{a,e\}$和$\{a,f\}$，删除$\{i,j\}$并且添加$\{g,i\}$和$\{g,j\}$。因此，彼得森图不是可平面的。■

练习

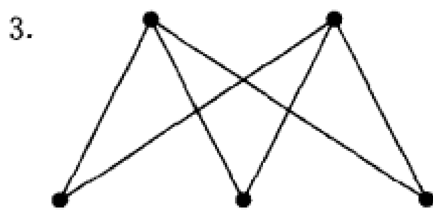
1. 五座房屋能否不带连接交叉地与两种设施相连接？

在练习 2～4 里，不带任何交叉地画出给定的可平面图。

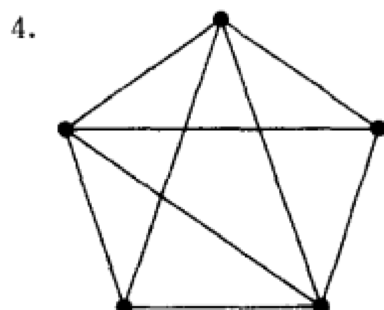
2.



3.

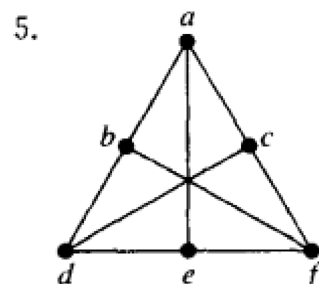


4.

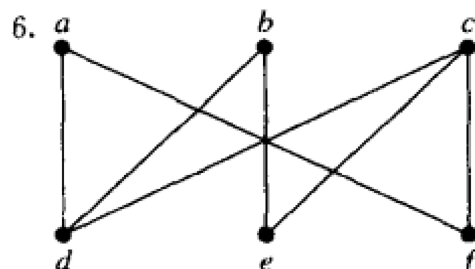


在练习 5～9 里，确定所给的图是否可平面。若是可平面的，则画出它使得没有边交叉。

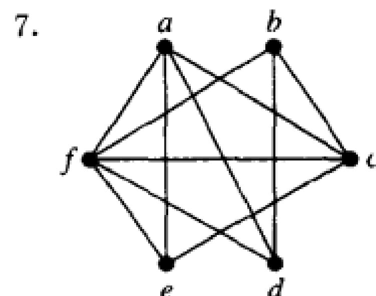
5.



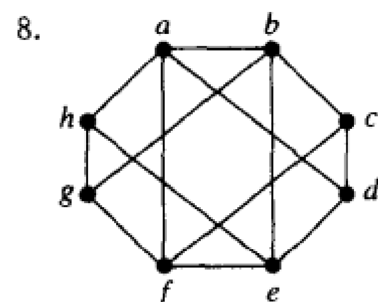
6.



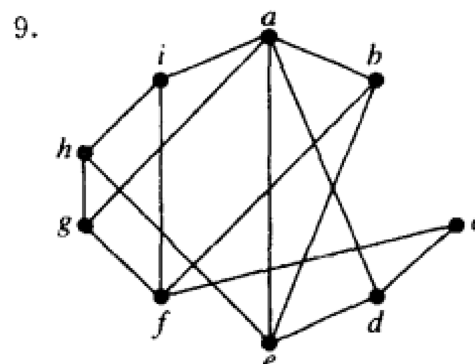
7.



8.



9.

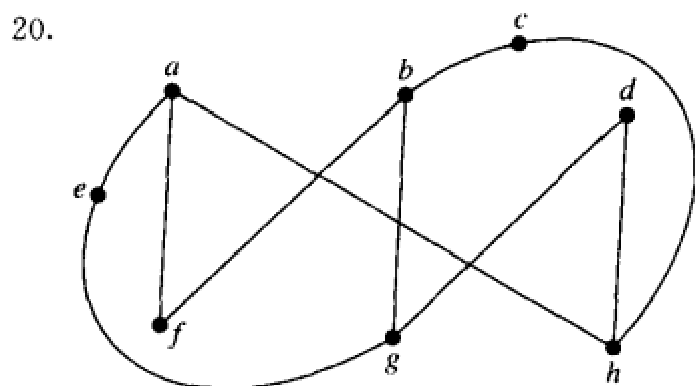


10. 完成例 3 里的论证。
11. 用类似于在例 3 里给出的论证来证明：K_5 是非可平面的。
12. 假定一个连通可平面图有 8 个顶点，每个顶点的度都为 3。这个图的可平面表示把平面分割成多少个面？
13. 假定一个连通可平面图有 6 个顶点，每个顶点的度都为 4。这个图的可平面表示把平面分割成多少个面？
14. 假定一个连通可平面图有 30 条边。若这个图的可平面表示把平面分割成 20 个面，则这个图有多少个顶点？
15. 证明推论 3。
16. 假定一个连通偶可平面图有 e 条边和 v 个顶点。证明：若 $v\geqslant 3$ 则 $e\leqslant 2v-4$。
*17. 假定一个带有 e 条边和 v 个顶点的连通可平面简单图不包含长度为 4 或更短的回路。证明：若 $v\geqslant 4$ 则 $e\leqslant (5/3)v-(10/3)$。
18. 假定一个可平面图有 k 个连通分支、e 条边和 v 个顶点。另外假定这个图的可平面表示把平面分割成 r 个面。求用 e，v 和 k 所表示的 r 的公式。
19. 下面的哪些非可平面图具有这样的性质：删除任何一个顶点以及与这个顶点关联的所有边就产生一个可平面图？

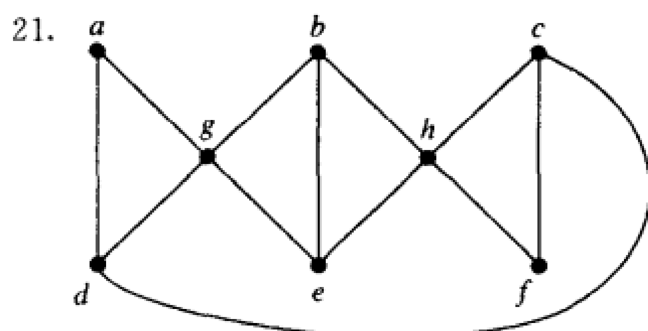
    a) K_5　　b) K_6　　c) $K_{3,3}$　　d) $K_{3,4}$

在练习 20～22 里，确定给定的图是否同胚于 $K_{3,3}$。

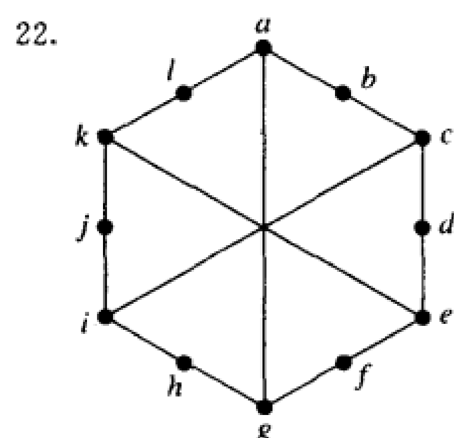
20.



21.

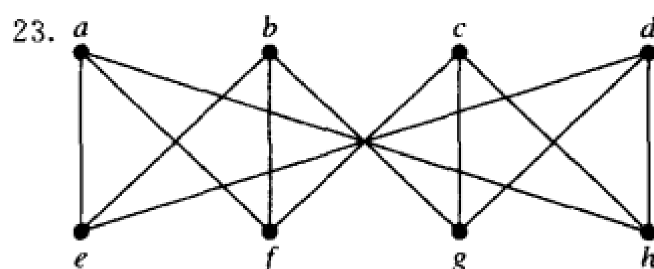


22.

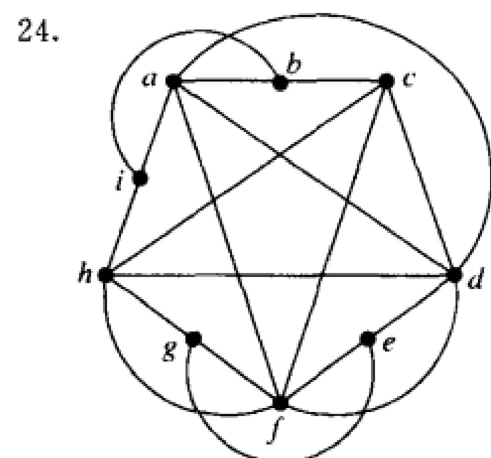


在练习 23～25 里，用库拉图斯基定理来确定所给的图是不是可平面的。

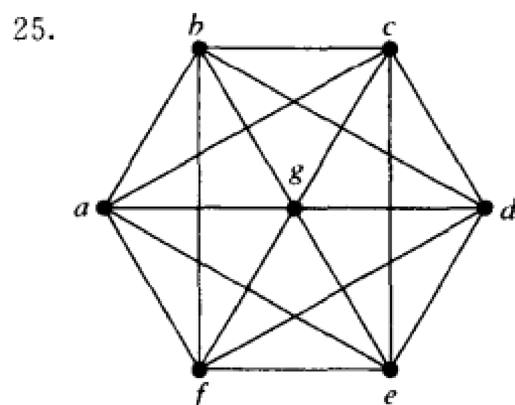
23.



24.



25.



一个简单图的交叉数是指，当在平面上画出这个图，其中不允许任何三条表示边的弧线在同一个点交叉时，交叉的最少次数。

26. 证明：$K_{3,3}$的交叉数为 1。

**27. 求下面每个非可平面图的交叉数。

a) K_5　　b) K_6　　c) K_7

d) $K_{3,4}$　　e) $K_{4,4}$　　f) $K_{5,5}$

*28. 求彼得森图的交叉数。

**29. 证明：若 m 和 n 都是偶正整数，则 $K_{m,n}$的交叉数小于或等于 $mn(m-2)(n-2)/16$。(提示：沿着 x 轴放置 m 个顶点，使得它们间距相等并且关于原点对称，再沿着 y 轴放置 n 个顶点，使得它们间距相等并且关于原点对称。现在连接 x 轴上 m 个顶点中的每一个与 y 轴上 n 个顶点中的每一个，并且计算交叉数。)

简单图 G 的厚度是指，以 G 作为它们的并图的 G 的可平面子图的最小个数。

30. 证明：$K_{3,3}$的厚度为 2。

*31. 求练习 27 里的图的厚度。

32. 证明：若 G 是一个带有 v 个顶点和 e 条边的连通简单图，则 G 的厚度至少为$\lceil e/(3v-6)\rceil$。

*33. 利用练习 32 来证明：每当 n 是正整数时，K_n 的厚度就至少为$\lfloor (n+7)/6\rfloor$。

34. 证明：若 G 是一个带有 v 个顶点和 e 条边并且没有长度为 3 的回路的连通简单图，则 G 的厚度至少为$\lceil e/(2v-4)\rceil$。

35. 利用练习34来证明：每当 m 和 n 都是正整数时，$K_{m,n}$ 的厚度就至少是 $\lceil mn(2m+2n-4)\rceil$。

*36. 在一个环面(原文 torus，指自行车轮胎或救生圈这类形状——译者注)上画出 K_5，使得没有边交叉。

*37. 在一个环面上画出 $K_{3,3}$，使得没有边交叉。

8.8 图着色

8.8.1 引言

在图论中，有许多与地图区域(比如世界各部分的地图)着色有关的理论成果。当为一幅地图着色时[⊖]，具有公共边界的两个区域在传统上指定为不同的颜色。一种确保两个相邻的区域永远没有相同的颜色的方法是对每个区域都使用不同的颜色。不过，这种方法效率不高，而且在具有许多区域的地图上，可能难以区分相似的颜色。另一种方法是，应当尽可能地使用少数几种颜色。考虑这样的问题：确定可以用来给一幅地图着色的颜色的最小数目，使得相邻的区域永远没有相同的颜色。例如，对图8-80左侧地图来说，四色是足够的，但是三色就不够(读者应当验证这一点)。对图8-80右侧地图来说，三色是足够的(但是二色就不够)。

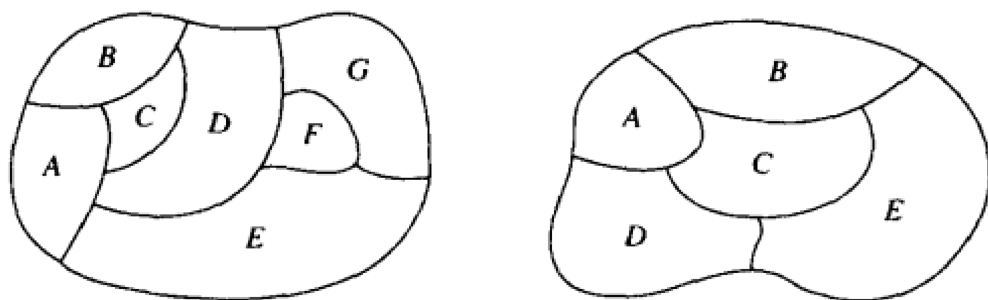


图8-80 两幅地图

平面里的每幅地图都可以表示成一个图。为了建立这样的对应关系，地图的每个区域都表示成一个顶点。若两个顶点所表示的区域具有公共边界，则用边连接这两个顶点。只相交于一个点的两个区域不算是相邻的。这样所得到的图称为这个地图的对偶图。根据地图的对偶图的构造方式，显然在平面里的任何地图都具有可平面的对偶图。图8-81显示对应于图8-80所示地图的对偶图。

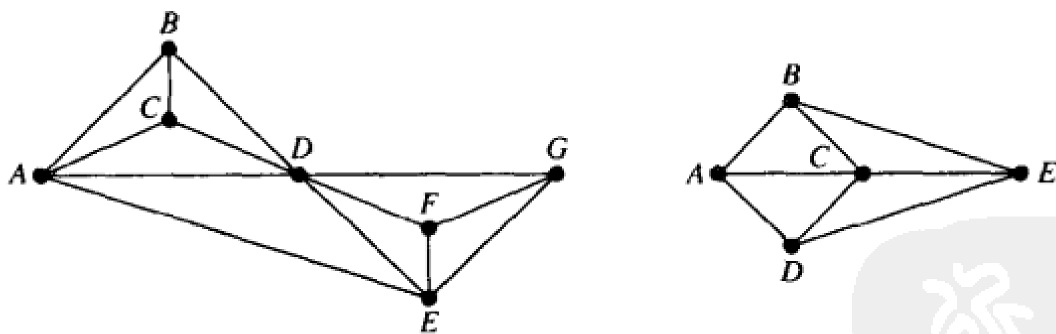


图8-81 图8-80地图的对偶图

给地图的区域着色的问题等价于这样的问题：给对偶图的顶点着色，使得在对偶图里没有两个相邻的顶点具有相同的颜色。给出下面的定义。

定义1 简单图的着色是对该图的每个顶点都指定一种颜色，使得没有两个相邻的顶点颜色相同。

⊖ 假定地图里所有区域都是连通的。这样就消除了像密歇根这样的地理状况所引起的问题，因为密歇根州是由两个分离的半岛组成的。

通过对每个顶点都指定一种不同的颜色，就可以着色一个图。不过，对大多数的图来说，可以找到所用颜色数少于图中顶点数的着色。什么是所需要的最少颜色数？

定义 2 图的色数是着色这个图所需要的最少颜色数。

注意，求平面图的色数等于是求平面地图着色所需要的最少颜色数，使得没有两个相邻的区域指定为相同的颜色。这个问题已经研究了 100 多年。数学里最著名的定理之一给出了它的答案。

定理 1 四色定理。平面图的色数不超过四。

四色定理最早是作为猜想在 19 世纪 50 年代提出的[㊀]。美国数学家肯尼思·阿佩尔和沃尔夫冈·黑肯最终在 1976 年证明了它。在 1976 年之前，发表过许多不正确的证明，其中的错误常常难以发现。另外，还尝试过画出需要超过四色的地图来构造反例，而这样做是无效的。(证明五色定理就没有这样困难，参见练习 35。)

也许迄今为止在数学里最有名的错误证明就是伦敦的律师和业余的数学家艾尔弗雷德·肯普[㊁]在 1879 年所发表的四色定理证明。数学家们一直认为他的证明是正确的，直到 1890 年珀西·希伍德发现了一处错误，才发现肯普的论证是不完全的。不过，事实证明，肯普的推理路线是阿佩尔和黑肯所给出的成功证明的基础。他们的证明依赖于计算机所完成的对各种情形的仔细分析。他们证明，若四色定理为假，则存在一个反例，它是大约 2000 种不同类型中的一种，然后他们证明这些类型都没有导致反例。在他们的证明中使用了 1000 多个小时的机时。计算机在证明过程中起到如此重要的作用，由此引发了广泛的争论。例如，在计算机程序里有没有导致不正确结果的错误？假如论证依赖于或许不可靠的计算机得出，那么它是不是真正的证明？[㊂]

注意，四色定理只适用于平面图。在例 2 里将证明非可平面图可以有任意大的色数。

证明一个图的色数为 n 需要做两件事。首先必须证明：用 n 种颜色可以着色这个图。构造出这样的着色就可以完成这件事。其次证明：用少于 n 种颜色不能着色这个图。例 1 至例 4 说明如何求出色数。

例 1 图 8-82 所示的图 G 和 H 的色数是什么？

解 图 G 的色数至少为 3，因为顶点 a，b 和 c 必须为不同的颜色。为了看出是否可以用三种颜色来对 G 着色，指定 a 为红，b 为蓝，c 为绿。于是，可以(而且必须)令 d 为红，因为它与 b 和 c 相邻。另外，可以(而且必须)令 e 为绿，因为它只与红色和蓝色顶点相邻，可以(而且必须)令 f 为蓝，因为它只与红色和绿色顶点相邻。最后，可以(而且必须)令 g 为红，因为它只与蓝色和绿色的顶点相邻。这样就产生出恰好使用三种颜色的 G 的着色，如图 8-83 所示。

㊀ *历史注解*。在 1852 年，德·摩根从前的一个学生弗朗西斯·古特利注意到，用四种颜色可以给英格兰的郡着色，使得没有相邻的郡颜色相同。在此基础上，他猜想四色定理为真。弗朗西斯把这个问题告诉他的弟弟弗雷德里克，弗雷德里克当时是德·摩根的学生。弗雷德里克就哥哥的猜想询问了他的老师德·摩根。德·摩根对这个问题极其感兴趣，并且向数学界公布了它。事实上，在德·摩根给威廉·罗万·哈密顿爵士的信中第一次在书面上提到这个猜想。虽然德·摩根认为哈密顿可能对这个问题也感兴趣，但事实完全不是这样，因为它与四元数毫无关系。

㊁ **艾尔弗雷德·布雷·肯普(Alfred Bray Kempe, 1849—1922)** 肯普是一位律师和教会法规的主要权威。不过，在剑桥大学学习数学之后，他依然保持着对数学的兴趣，并且在以后的生活中对数学研究投入了相当多的时间。肯普对动力学和逻辑学做出过贡献，动力学是处理运动的数学分支。不过，肯普被人记住，主要还是因为他对四色定理的错误证明。

㊂ *历史注解*。虽然罗布森、桑得尔斯、西摩尔以及托马斯在 1996 年找到了四色定理的简化证明，把证明的计算部分减少到检查 633 种格局，但是仍然没有找到不依赖于大量计算的证明。

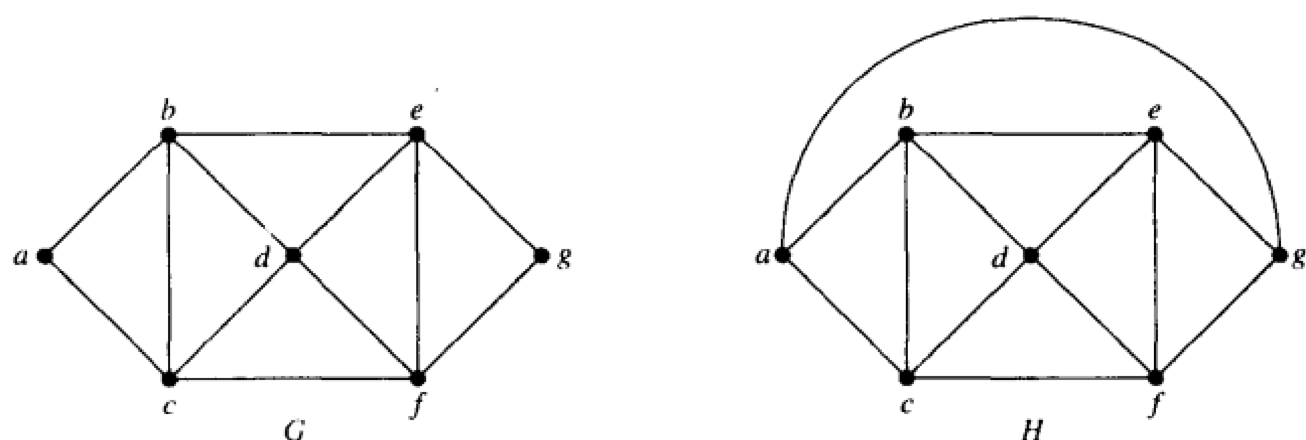


图 8-82 简单图 G 和 H

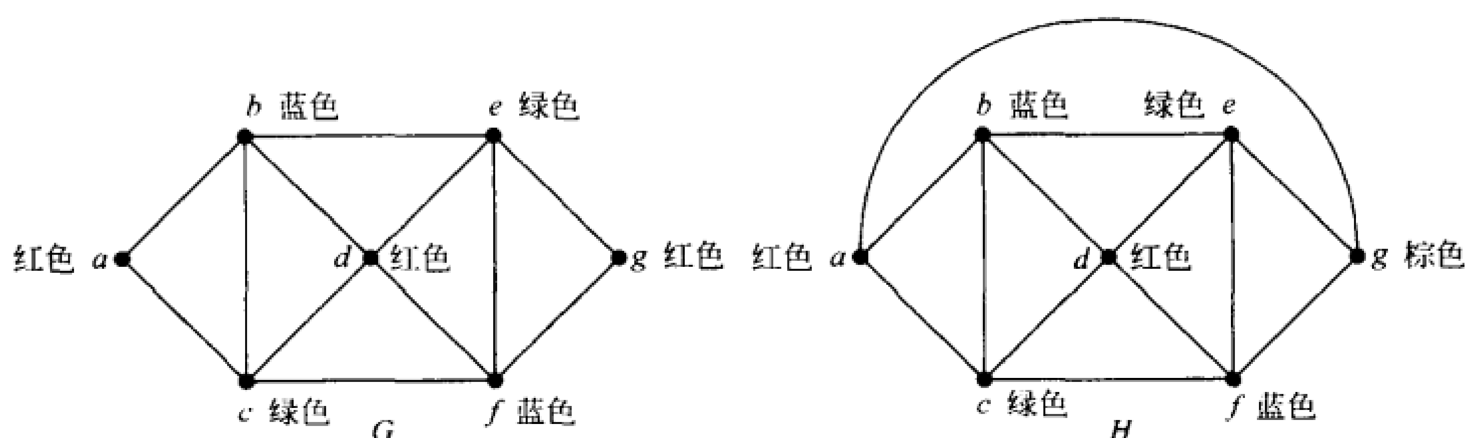


图 8-83 图 G 和 H 的着色

图 H 是由图 G 和连接 a 与 g 的一条边所组成的。用三种颜色来着色 H 的任何尝试都必须遵循与着色 G 所用的同样的推理，不同之处是在最后阶段当除了 g 以外的所有顶点都已经着色时，因为 g 与红色、蓝色和绿色顶点（在 H 里）相邻，所以需要使用第四种颜色，比如棕色。因此，H 的色数为四。图 8-83 右图显示 H 的一种着色。■

例 2 K_n 的色数是什么？

解 通过给每个顶点指定一种不同的颜色，用 n 种颜色可以构造 K_n 的着色。使用的颜色能否更少一些？答案是不能。没有两个顶点可以指定相同的颜色，因为这个图的每两个顶点都是相邻的。因此，K_n的色数$=n$。（回忆一下，当 $n\geqslant 5$ 时 K_n不是可平面图，所以这个结果与四色定理并不矛盾。）在图 8-84 里显示使用五种颜色的 K_5 的着色。■

例 3 完全偶图 $K_{m,n}$的色数是什么？其中 m 和 n 都是正整数。

解 需要的颜色数似乎依赖于 m 和 n。不过，仅仅需要两种颜色。m 个顶点为一种颜色，n 个顶点为另外一种颜色。因为边都只能连接 m 个顶点中的一个顶点与 n 个顶点中的一个顶点，所以没有相邻的顶点具有相同的颜色。在图 8-85 里显示带有两种颜色的 $K_{3,4}$的着色。■

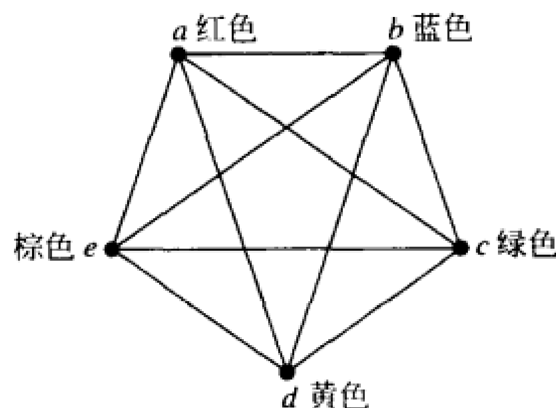


图 8-84 K_5 的着色

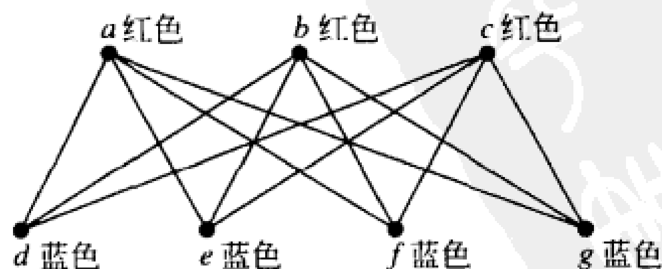


图 8-85 $K_{3,4}$的着色

每个连通偶简单图都具有色数 2 或 1，因为在例 3 里使用的推理适用于任何这样的图。反过来，每个色数为 2 的图都是偶图(见本节末的练习 25 和练习 26)。

例 4 图 C_n 的色数是什么？(回忆一下，C_n是带有 n 个顶点的圈图。)

解 首先考虑一些个别情形。首先，设 $n=6$。挑选一个顶点并且把它着色成红色。在图 8-86所示的 C_6 的平面画法里顺时针前进。必须给到达的下一个顶点指定第二种颜色，比如蓝色。以顺时针方向继续下去，可以令第三个顶点为红色，第四个顶点为蓝色，第五个顶点为红色。最后，令第六个顶点为蓝色，它与第一个顶点是相邻的。因此，C_6 的色数为 2。图 8-86 显示这里构造的着色。

其次，设 $n=5$ 并且考虑 C_5。挑选一个顶点并且令它为红色。顺时针前进，必须给到达的下一个顶点指定第二种颜色，比如蓝色。以顺时针方向继续下去，可以令第三个顶点为红色，第四个顶点为蓝色。第五个顶点既不能为红色也不能为蓝色，因为它与第四个顶点和第一个顶点都相邻。所以，对这个顶点就需要第三种颜色。注意，假如以逆时针方向对顶点着色，同样需要三种颜色。因此，C_5 的色数是 3。用 3 种颜色对 C_5 着色见图 8-86。

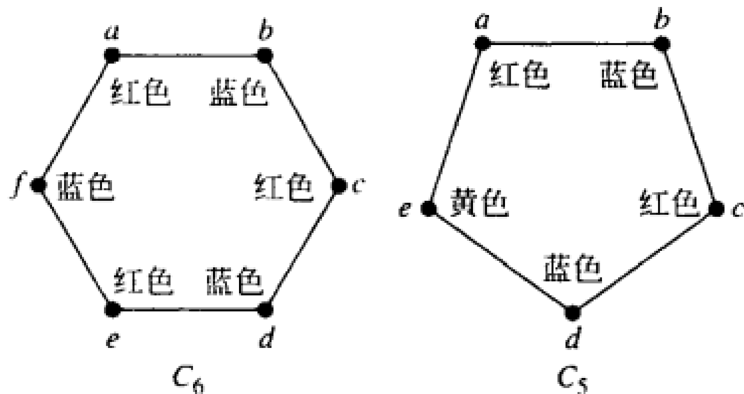


图 8-86 C_5 和 C_6 的着色

在一般情形里，当 n 是偶数时，对 C_n着色需要两种颜色。为了构造这样的着色，简单地挑选一个顶点并且令它为红色。然后(利用图的平面表示)以顺时针方向绕图前进，令第二个顶点为蓝色，第三个顶点为红色，依此类推。可以令第 n 个顶点为蓝色，因为与它相邻的两个顶点(即第 $n-1$ 个顶点和第一个顶点)都是红色。

当 n 是奇数而且 $n>1$ 时，C_n的色数为 3。为了看出这一点，挑选一个初始顶点。为了只用两种颜色，当以顺时针方向遍历这个图时，必须交替使用颜色。不过，所到达的第 n 个顶点与带不同颜色的两个顶点(第一个顶点和第 $n-1$ 个顶点)相邻。因此，必须使用第三种颜色。 ■



已知最好的求图的色数的算法(对图的顶点数来说)具有指数的最坏情形时间复杂性。即使求图的色数的近似值也是很难的。已经证明，假如存在具有多项式最坏情形时间复杂性的、可以达到 2 倍地近似图的色数的算法(即构造出一个不超过图的色数的两倍的界限)，那么也存在具有多项式最坏情形时间复杂性的求图的色数的算法。

8.8.2 图着色的应用

图着色在与调度和分配有关的问题中具有多种应用。(注意，由于不知道图着色的有效算法，所以这并不能得出调度和分配的有效算法。)在这里将给出这样的应用的例子。第一个应用是用来安排期末考试。

例 5 **安排期末考试** 如何安排一所大学里的期末考试，使得没有学生同时要考两门试？

解 这样的安排问题可以用图模型来解决，用顶点表示科目，若有学生要考两门试，则在表示考试科目的两个顶点之间有边。用不同颜色来表示期末考试的每个时间段。考试的安排就对应于所关联的图的着色。

例如，假定要安排七门期末考试。假定科目从 1 到 7 编号。假定下列各对科目的考试有学生要都参加：1 和 2，1 和 3，1 和 4，1 和 7，2 和 3，2 和 4，2 和 5，2 和 7，3 和 4，3 和 6，3 和 7，4 和 5，4 和 6，5 和 6，5 和 7，以及 6 和 7。在图 8-87 里显示这组科目所关联的图。一种安排就是由这个图的一种着色来组成的。

因为这个图的色数为4(读者应当验证这一点)，所以需要四个时间段。图8-88显示使用了四种颜色的这个图的着色以及所关联的调度。 ■

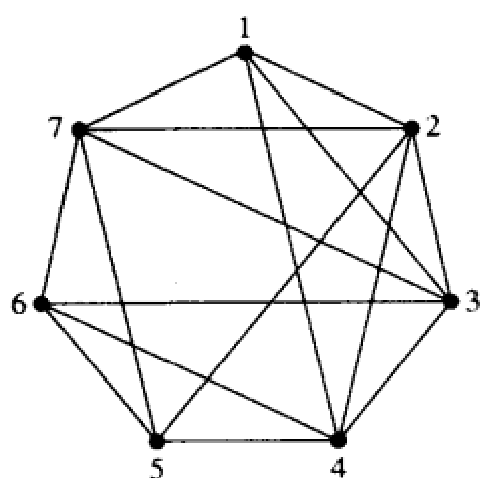
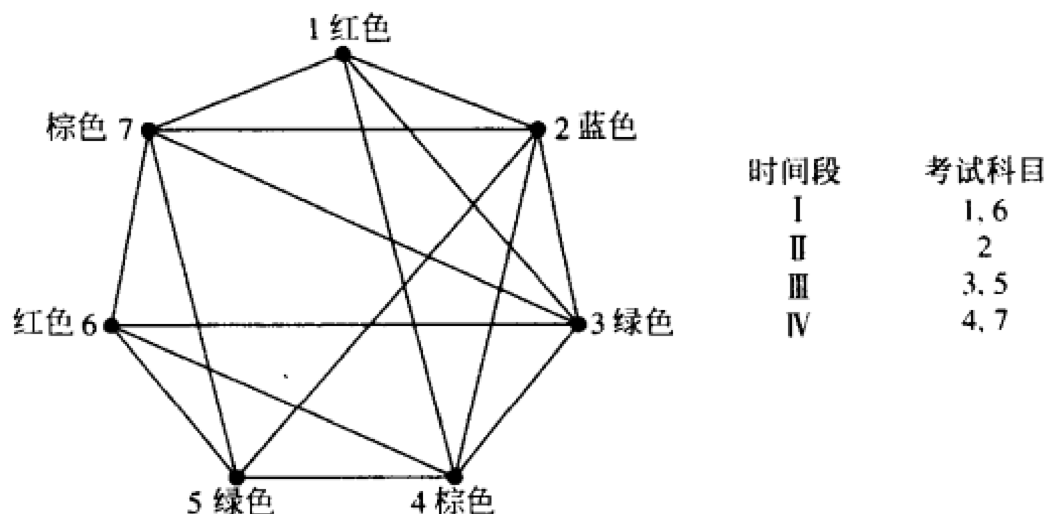


图 8-87 表示期末考试安排的图



时间段	考试科目
Ⅰ	1, 6
Ⅱ	2
Ⅲ	3, 5
Ⅳ	4, 7

图 8-88 用着色来安排期末考试

现在考虑对电视频道的分配。

例 6 频率分配 把频道2到13分配给在北美洲的电视台，要避免150英里之内的两家电视台在相同频道上播出。如何用图为频道分配建模？

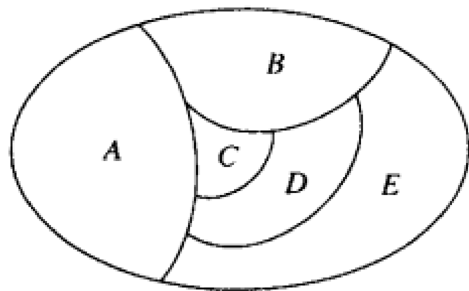
解 这样构造一个图：给每个电视台指定一个顶点。若两个电视台彼此位于150英里以内，则用边连接这两个顶点。频道指派就对应于这个图的着色，其中每种颜色表示一个不同的频道。 ■

例 7 变址寄存器 在有效的编译器里，当把频繁地使用的变量暂时保存在中央处理单元的变址寄存器里，而不是保存在常规内存里时，可以加速循环的执行。对于给定的循环来说，需要多少个变址寄存器？可以用图着色模型来讨论这个问题。为了建立这个模型，设图的每个顶点表示循环里的一个变量。若在循环执行期间两个顶点所表示的变量必须同时保存在变址寄存器里，则在这两个顶点之间有边。所以，这个图的色数就给出所需要的变址寄存器数，因为当表示变量的顶点在图中相邻时，就必须给这些变量分配不同的寄存器。 ■

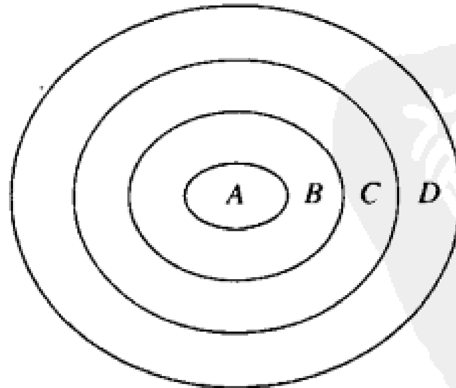
练习

在练习1～4里，构造所示地图的对偶图。然后求给这个地图着色使得相邻的两个区域都没有相同的颜色所需要的颜色数。

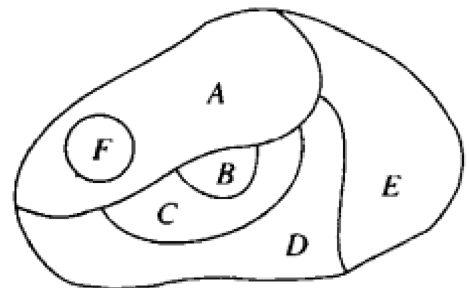
1.



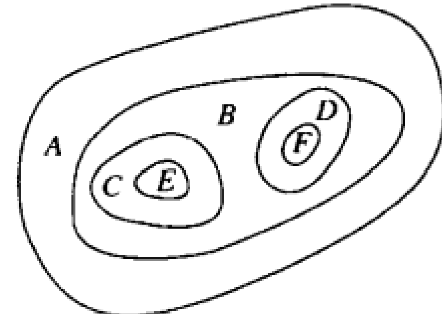
2.



3.

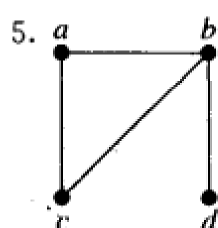


4.

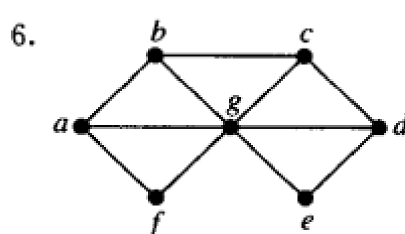


在练习 5～11 里，求给定的图的色数。

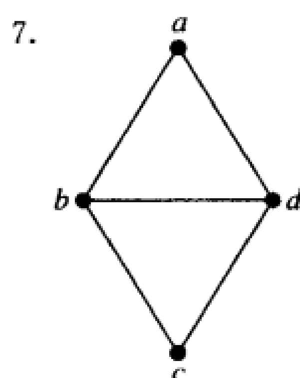
5.



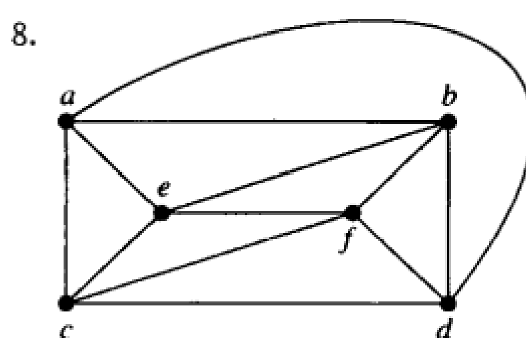
6.



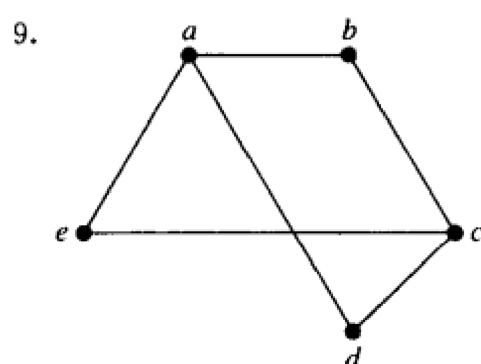
7.



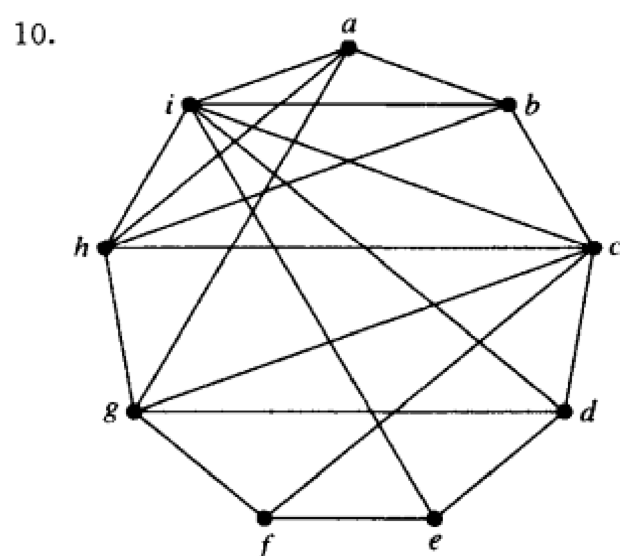
8.



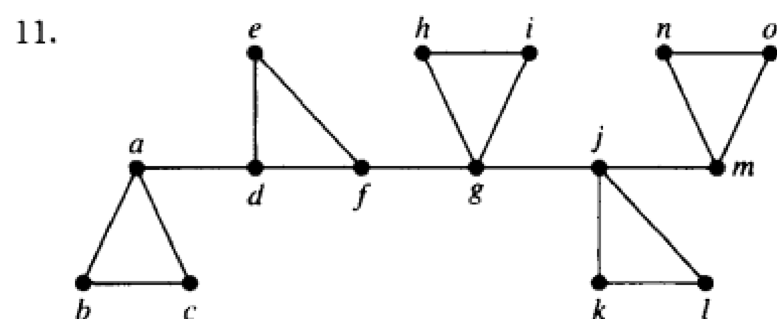
9.



10.



11.



12. 对练习 5～11 里的图，判定是否可能通过删除单个顶点和与它关联的所有边，来减少色数。

13. 哪些图具有色数 1？

14. 为美国地图着色所需要的最少颜色数是什么？不要考虑只在一个尖角处相触的相邻州。假定密歇根是一个区域。把表示阿拉斯加和夏威夷的顶点当作孤立点。

15. W_n 的色数是什么？

16. 证明：具有包含奇数个顶点的回路的简单图不能用两种颜色来着色。

17. 假定除了科目 Math 115 与 CS 473，Math 116 与 CS 473，Math 195 与 CS 101，Math 195 与 CS 102，Math 115 与 Math 116，Math 115 与 Math 185，Math 185 与 Math 195 以外，其他每种组合里都有学生要参加组合里所有科目的考试，请使用最少个数的不同时间段，来为 Math 115，Math 116，Math 185，Math 195，CS 101，CS 102，CS 273 和 CS 473 安排期末考试日程表。

18. 假定当两家电视台相距 150 英里以内时，它们就不能使用相同的频道，那么对相对距离如下表所示的六家电视台来说，需要多少个不同的频道？

	1	2	3	4	5	6
1	—	85	175	200	50	100
2	85	—	125	175	100	160
3	175	125	—	100	200	250
4	200	175	100	—	210	220
5	50	100	200	210	—	100
6	100	160	250	220	100	—

19. 数学系有六个委员会，都是每月开一次会。假定委员会是 C_1＝{阿林豪斯，布兰德，沙斯拉夫斯基}，C_2＝{布兰德，李，罗森}，C_3＝{阿林豪斯，罗森，沙斯拉夫斯基}，C_4＝{李，罗森，沙斯拉夫斯基}，C_5＝{阿林豪斯，布兰德}和 C_6＝{布兰德，罗森，沙斯拉夫斯基}，那么怎样安排才能确保没有人同时参加两个会议。
20. 动物园想建立自然居住地，在里面展出它的动物。然而，一些动物一有机会就会吃掉另外一些动物。如何用图模型和着色来确定所需要的不同居住地的数目，以及在这些居住地里的动物安排？

图的**边着色**是指对各边指定颜色，使得关联到相同顶点的边指定不同的颜色。图的**边色数**是在该图的边着色里可以使用的最少颜色数。

21. 求练习 5～11 里每个图的边色数。

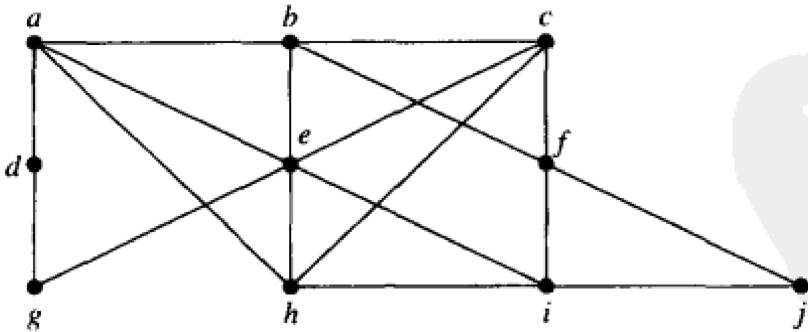
*22. 求下列图的边色数

a) K_n；　　b) $K_{m,n}$；　　c) C_n；　　d) W_n.

23. 七个变量出现在计算机程序的循环里。这些变量以及必须保存它们的计算步骤是：t：步骤 1 到 6；u：步骤 2；v：步骤 2 到 4；w：步骤 1，3 和 5；x：步骤 1 和 6；y：步骤 3 到 6；以及 z：步骤 4 和 5。在执行期间需要多少个不同的变址寄存器来保存这些变量？
24. 关于一个以 K_n 作为子图的图的色数能有些什么结论？
25. 证明：色数为 2 的简单图是偶图。
26. 证明：连通偶图的色数为 2。

下面的算法可以用来为简单图着色。首先，以度递减的顺序列出顶点 $v_1, v_2, \cdots, v_n$ 使得 $\deg(v_1)\geqslant\deg(v_2)\geqslant\cdots\geqslant\deg(v_n)$。把颜色 1 指定给 v_1 和在表中不与 v_1 相邻的下一个顶点(若存在一个这样的顶点)，并且继续指定给每一个在表中不与已经指定了颜色 1 的顶点相邻的顶点。然后把颜色 2 指定给表中还没有着色的第一个顶点。继续把颜色 2 指定给那些在表中还没有着色、而且不与指定了颜色 2 的顶点相邻的顶点。若还有未着色的顶点，则指定颜色 3 给表中还没有着色的第一个顶点，并且用颜色 3 继续对还没有着色、而且不与指定了颜色 3 的顶点相邻的那些顶点着色。继续这个过程直到所有顶点都着色为止。

27. 用这个算法构造下图的着色。

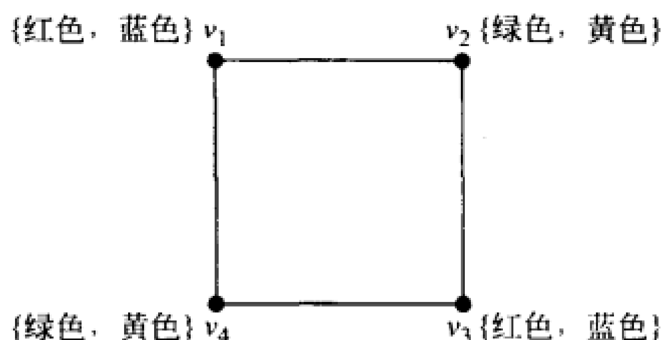


*28. 用伪代码描述这个着色算法。

*29. 证明：这个算法所产生的着色数可能比着色一个图所需的颜色数更多。

图 G 的 **k 重着色**是对 G 的顶点指定含有 k 种不同颜色的集合，使得相邻的顶点不具有相同的颜色。用 $\chi_k(G)$ 来表示使得 G 有用 n 种颜色的 k 重着色的最小正整数 n。例如，$\chi_2(C_4)=4$。为了看出这一点，注意如图所示，只用四种颜色，就可以对 C_4 的每个顶点指定两种颜色，使得两个相邻顶点不具有相同的颜色。

另外，少于四种颜色是不够的，因为顶点 v_1 和 v_2 每个都必须指定两种颜色，而且不能对 v_1 和 v_2 指定相同颜色。（关于 k 重着色的更多信息，见[MiRo90]。）



30. 求下列这些值：

a) $\chi_2(K_3)$　b) $\chi_2(K_4)$　c) $\chi_2(W_4)$　d) $\chi_2(C_5)$

e) $\chi_2(K_{3,4})$　f) $\chi_3(K_5)$　*g) $\chi_3(C_5)$　h) $\chi_3(K_{4,5})$

*31. 设 G 和 H 是图 8-82 所示的图。求

a) $\chi_2(G)$　b) $\chi_2(H)$　c) $\chi_3(G)$　d) $\chi_3(H)$

32. 若 G 是偶图而 k 是正整数，则 $\chi_k(G)$ 是什么？

33. 移动广播（或蜂窝）电话的频率是按地段分配的。每个地段分配一组该地段里的车辆所使用的频率。在产生干扰问题的地段里不能使用相同频率。解释一下如何用 k 重着色来对一个地区里的每个移动广播地段分配 k 种频率。

*34. 证明：每个可平面图 G 都可以用不超过 6 种颜色来着色。（提示：对图的顶点数用数学归纳法。应用 8.7 节推论 2 来求出顶点 v 满足 $\deg(v)\leqslant 5$。考虑删除 v 及其关联的所有边所获得的 G 的子图。）

**35. 证明：每个可平面图 G 都可以用不超过 5 种颜色来着色。（提示：利用练习 34 的提示。）

关键术语和结果

术语

无向边：与集合 $\{u, v\}$ 关联的边，其中 u 和 v 都是顶点。

有向边：与有序对 (u, v) 关联的边，其中 u 和 v 都是顶点。

多重边：连接同样一对顶点的不同的边。

环：连接一个顶点与它自身的边。

无向图：一组顶点以及连接这些顶点的一组无向边。

简单图：没有多重边和环的无向图。

多重图：可能包含多重边但不包含环的无向图。

伪图：可能包含多重边和环的无向图。

有向图：一组顶点以及连接这些顶点的一组有向边。

有向多重图：可能包含多重有向边的带有向边的图。

相邻：若在两个顶点之间有边则它们是相邻的。

关联：若一个顶点是一条边的端点则那条边关联那个顶点。

$\deg(v)$（无向图里顶点 v 的度）：与 v 关联的边的数目。

$\deg^-(v)$（带有向边的图里顶点 v 的入度）：以 v 作为终点的边的数目。

$\deg^+(v)$（带有向边的图里顶点 v 的出度）：以 v 作为起点的边的数目。

带有向边的图的底图：通过忽略边的方向所获得的无向图。

K_n（n 个顶点的完全图）：带 n 个顶点的无向图，其中每对顶点都用一条边连接。

偶图：顶点集划分成两个子集合 V_1 和 V_2 的图，使得每条边都连接 V_1 里的顶点和 V_2 里的顶点。

$K_{m,n}$**（完全偶图）**：顶点集划分成 m 个元素的子集和 n 个元素的子集，使得两个顶点被一条边所连接，当且仅当一个顶点属于第一个子集而另外一个顶点属于第二个子集。

C_n**（大小为 n 的圈图）**，$n \geqslant 3$：带有 n 个顶点 v_1，v_2，…，v_n 和边 $\{v_1, v_2\}$，$\{v_2, v_3\}$，…，$\{v_{n-1}, v_n\}$，$\{v_n, v_1\}$ 的图。

W_n**（大小为 n 的轮图）**，$n \geqslant 3$：通过向 C_n 添加一个顶点以及从这个顶点到 C_n 里原来每个顶点的一条边所获得的图。

Q_n**（n 立方体图）**，$n \geqslant 1$：用 2^n 个长度为 n 的位串作为顶点，边连接恰好相差一位的每对位串的图。

孤立点：0 度顶点。

悬挂点：1 度顶点。

正则图：所有顶点都有相同的度的图。

图 $G=(V, E)$ 的子图：图 (W, F)，其中 W 是 V 的子图而 F 是 E 的子图。

$G_1 \cup G_2$**（G_1 与 G_2 的并图）**：图 $(V_1 \cup V_2, E_1 \cup E_2)$，其中 $G_1=(V_1, E_1)$ 而 $G_2=(V_2, E_2)$。

邻接矩阵：利用顶点的相邻关系来表示图的矩阵。

关联矩阵：利用边与顶点的关联关系来表示图的矩阵。

同构的简单图：对简单图 $G_1=(V_1, E_1)$ 和简单图 $G_2=(V_2, E_2)$ 来说，若存在从 V_1 到 V_2 的一一对应 f，使得对所有属于 V_1 的 v_1 和 v_2 来说，$\{f(v_1), f(v_2)\} \in E_2$ 当且仅当 $\{v_1, v_2\} \in E_1$，则 G_1 与 G_2 是同构的。

同构不变量：同构的图都有或都没有的性质。

无向图里从 u 到 v 的通路：一条或多条边的序列 e_1，e_2，…，e_n，其中对 $i=0, 1, \cdots, n$ 来说，e_i 关联着 $\{x_i, x_{i+1}\}$，其中 $x_0=u$ 而 $x_{n+1}=v$。

有向图里从 u 到 v 的通路：一条或多条边的序列 e_1，e_2，…，e_n，其中对 $i=0, 1, \cdots, n$ 来说，e_i 关联着 (x_i, x_{i+1})，其中 $x_0=u$ 而 $x_{n+1}=v$。

简单通路：不多次包含一条边的通路。

回路：在相同顶点处开始与结束的通路，通路长度 $n \geqslant 1$。

连通图：在图中每对顶点之间都有通路的无向图。

图 G 的连通分支：图 G 的最大连通子图。

强连通有向图：从每个顶点到每个顶点都存在有向通路的有向图。

有向图 G 的强连通分支：G 的最大强连通子图。

欧拉回路：恰好包含图的每条边一次的回路。

欧拉通路：恰好包含图的每条边一次的通路。

哈密顿通路：简单图 $G=(V, E)$ 里的通路 x_0，x_1，…，x_n，使得 $\{x_0, x_1, \cdots, x_n\}=V$ 而且对 $0 \leqslant i < j \leqslant n$ 来说有 $x_i \neq x_j$。

哈密顿回路：简单图里的回路 x_0，x_1，…，x_n，x_0，使得 x_0，x_1，…，x_n 是哈密顿通路。

带权图：为各边指定数字的图。

最短通路问题：确定带权图里的通路以使得这条通路里的边的权之和在所规定的顶点之间的所有通路上是最小值这样的问题。

旅行商问题：求访问图的每个顶点恰好一次的、总长度最短的回路的问题。

可平面图：可以画在平面上而没有边交叉的图。

可平面图的平面表示面：该图的平面表示把平面所分割成的区域。

初等细分：删除无向图的边$\{u, v\}$而且添加新顶点 w 以及边$\{u, w\}$和边$\{w, v\}$。

同胚：若两个无向图是从同一个无向图通过一系列初等细分来获得的，则它们同胚。

图着色：给图的顶点指定颜色，使得相邻的两个顶点没有相同的颜色。

色数：在图的着色里所需要的最少颜色数。

结果

在连通多重图里存在欧拉回路当且仅当每个顶点都有偶数度。

在连通多重图里存在欧拉通路当且仅当至多两个顶点有奇数度。

迪克斯特拉算法：在带权图里求出两个顶点之间最短通路的过程(见 8.6 节)。

欧拉公式：$r=e-v+2$，其中 r，e 和 v 分别是可平面图的平面表示的面数、边数和顶点数。

库拉图斯基定理：图是不可平面的当且仅当它包含同胚于 $K_{3,3}$ 或 K_5 的子图。(其证明超出本书范围。)

四色定理：每个平面图都可以用不超过四种颜色来着色。(其证明远远超出本书范围!)

复习题

1. a)定义：简单图，多重图，伪图，有向图，有向多重图。
   b)用例子说明：如何用 a)里每种类型的图来建模。例如，解释一下如何为计算机网络或飞行航线的不同方面来建模。
2. 给出如何用图建模的至少四个例子。
3. 在无向图里，顶点度数之和与该图里边数之间的关系是什么？解释这个关系为什么成立。
4. 为什么在无向图里奇数度顶点的个数必然是偶数？
5. 在有向图里顶点的入度之和与出度之和之间的关系是什么？解释这个关系为什么成立。
6. 描述下列图族。
   a)K_n，在 n 个顶点上的完全图
   b)$K_{m,n}$，在 m 和 n 个顶点上的完全偶图
   c)C_n，带 n 个顶点的圈图
   d)W_n，大小为 n 的轮图
   e)Q_n，n 立方体
7. 在练习 6 的图族里，每个图有多少个顶点和多少条边？
8. a)什么是偶图？
   b)图 K_n，C_n 和 W_n 中哪些是偶图？
   c)你如何确定无向图是否为偶图？
9. a)描述用来表示图的三种不同方法。
   b)画出至少带 5 个顶点和 8 条边的简单图。说明如何用你在 a)里所描述的方法来表示它。
10. a)两个简单图是同构的是什么意思？
    b)对于简单图的同构来说，不变量是什么意思？给出至少五个这样的不变量的例子。
    c)给出带有相同的顶点数、边数和顶点度但不同构的两个简单图的例子。
    d)是否有一组已知的不变量可以用来有效地确定两个简单图是否同构？
11. a)图是连通的是什么意思？
    b)什么是图的连通分支？
12. a)解释一下如何用邻接矩阵来表示图。

b)如何用邻接矩阵来确定从图 G 的顶点集到图 H 的顶点集的函数是否是同构的?

c)如何用图的邻接矩阵来确定在图的两个顶点之间长度为 r 的通路数?其中 r 是正整数。

13. a)定义无向图里的欧拉回路和欧拉通路。

b)描述著名的哥尼斯堡七桥问题,并且解释一下如何利用欧拉回路来重新叙述它。

c)如何确定无向图是否具有欧拉通路?

d)如何确定无向图是否具有欧拉回路?

14. a)定义简单图里的哈密顿回路。

b)给出一些性质,这些性质蕴含着简单图没有哈密顿回路。

15. 给出至少两个可以通过求出带权图里最短通路来解决的问题的例子。

16. a)描述求在带权图两个顶点之间的最短通路的迪克斯特拉算法。

b)画出至少带 10 个顶点和 20 条边的带权图。用迪克斯特拉算法求出在图中你所选择的两个顶点之间的最短通路。

17. a)图是可平面的是什么意思?

b)给出不可平面的简单图的例子。

18. a)平面图的欧拉公式是什么?

b)如何用平面图的欧拉公式来证明简单图是不可平面的?

19. 叙述关于图的可平面性的库拉图斯基定理,并且解释一下它如何刻画了哪些图是可平面的。

20. a)定义图的色数。

b)当 n 是正整数时,图 K_n 的色数是什么?

c)当 n 是大于 2 的正整数时,图 C_n 的色数是什么?

d)当 m 和 n 都是正整数时,图 $K_{m,n}$ 的色数是什么?

21. 叙述四色定理。有没有不能用四种颜色来着色的图?

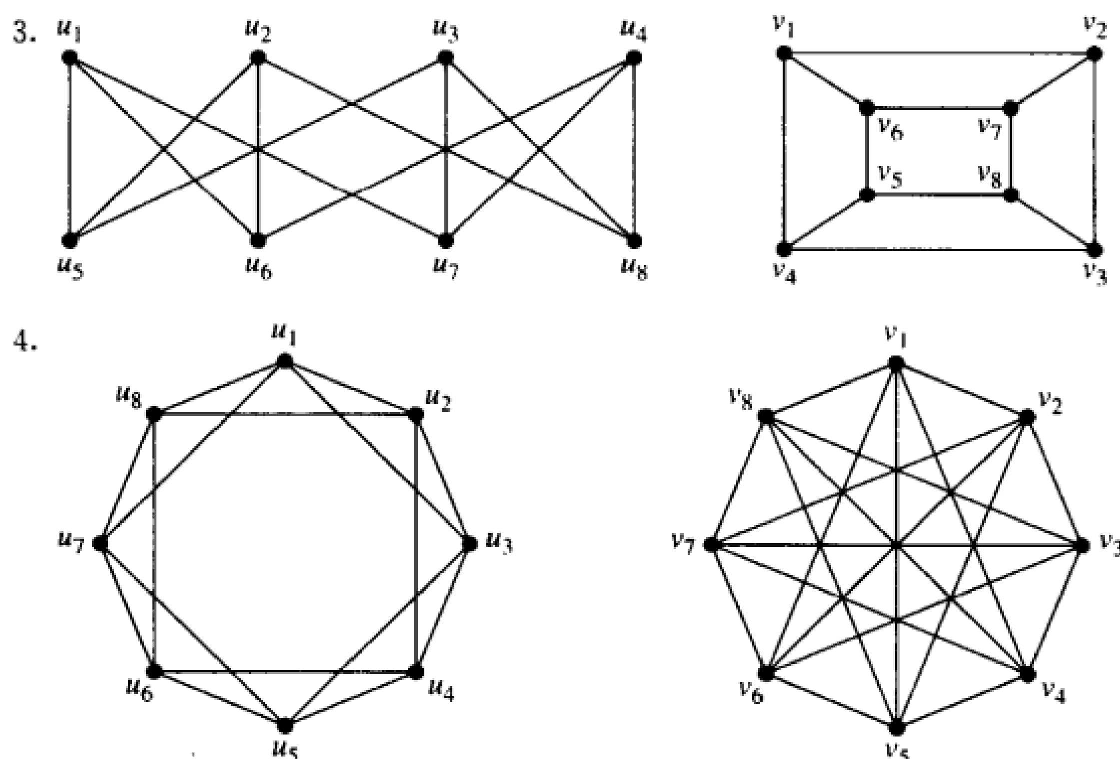
22. 解释一下在建模里可以如何使用图的着色。至少举两个不同的例子。

补充练习

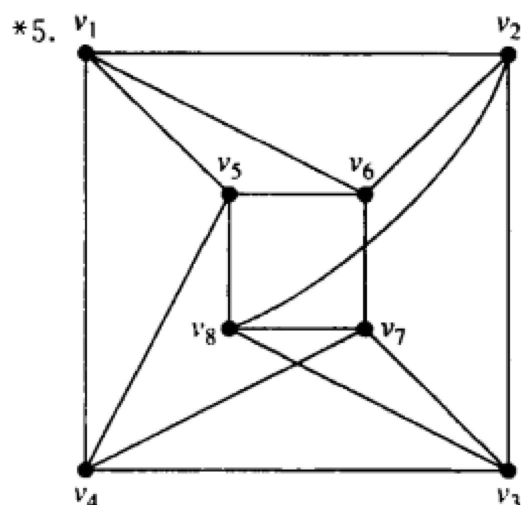
1. 一个带 100 个顶点的 50 正则图有多少条边?

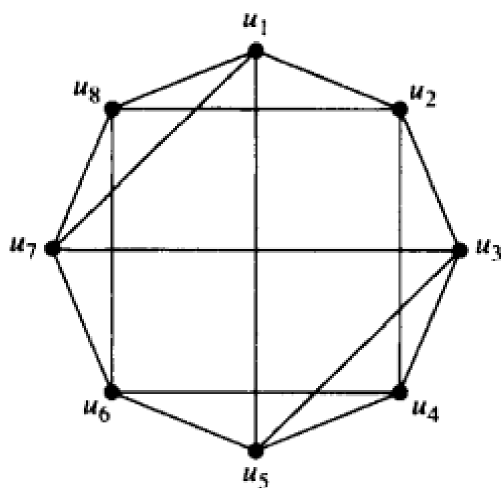
2. K_3 有多少种非同构的子图?

在练习 3~5 里,确定所给的成对的图是否同构。



*5.





完全 m 部图 $K_{n1,n2,\cdots,nm}$ 的顶点划分成 m 个子集合，各有 n_1，n_2，…，n_m 个元素，而且顶点相邻当且仅当它们属于这个划分的不同子集合。

6. 画出下列图。

a) $K_{1,2,3}$　　b) $K_{2,2,2}$　　c) $K_{1,2,2,3}$

*7. 完全 m 部图 $K_{n1,n2,\cdots,nm}$ 有多少个顶点和多少条边？

*8. a)证明或反驳：在至少有两个顶点的有穷简单图里，总是存在两个度数相同的顶点。

b)对有穷多重图做与(a)里同样的事。

设 $G=(V, E)$ 是简单图。顶点集合 V 的子集合 W 的导出子图是图 (W, F)，其中边集合 F 包含着 E 里的一条边当且仅当这条边的两个端点都属于 W。

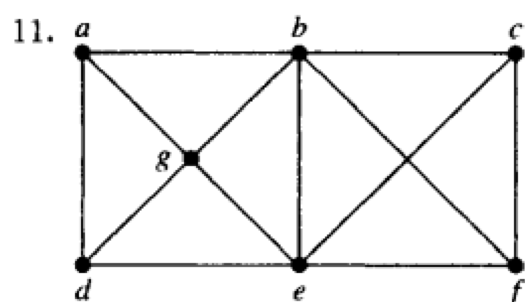
9. 考虑 8.4 节图 8-41 所示的图。求下列顶点的导出子图

a) $\{a, b, c\}$　　b) $\{a, e, g\}$　　c) $\{b, c, f, g, h\}$

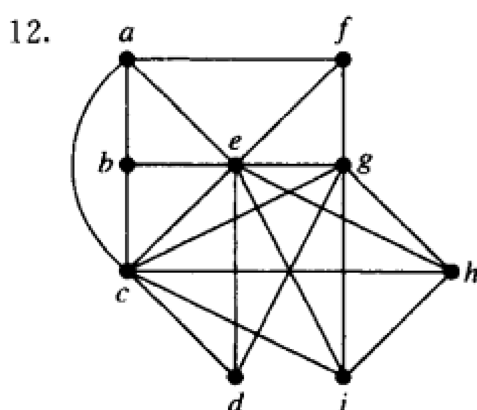
10. 设 n 是正整数。证明：K_n 的顶点集的非空子集的导出子图是完全图。

简单无向图里的团是一个完全子图，它不包含在任何更大的完全子图里。在练习 11～13 里，求所给的图的所有团。

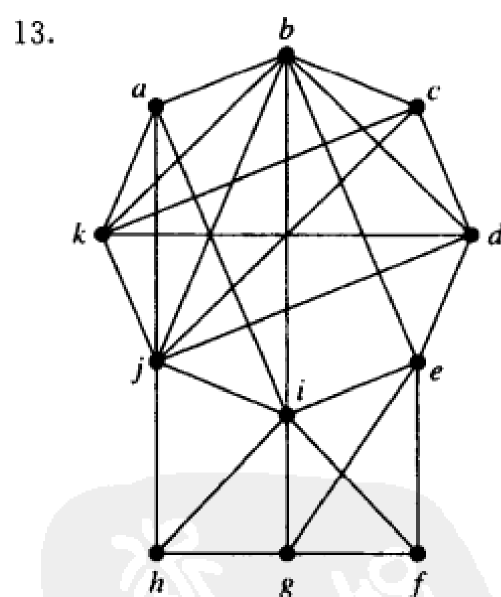
11.



12.

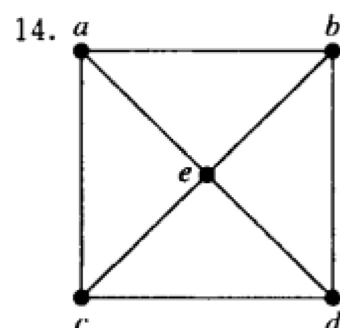


13.

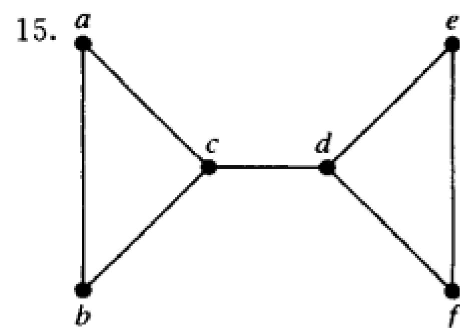


简单图里顶点的支配集是顶点的一个集合，使得其他每个顶点都与这个集合里至少一个顶点是相邻的。带最少顶点数的支配集称为最小支配集。在练习 14～16 里，求所给的图的最小支配集。

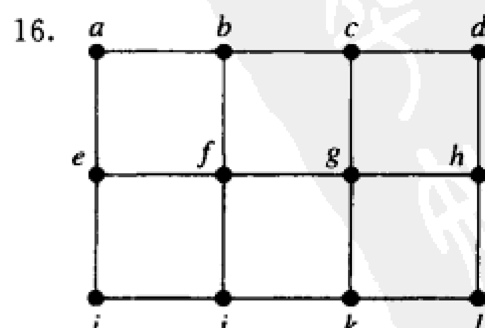
14.



15.



16.



简单图可用来确定在棋盘上控制整个棋盘的最少皇后数。一个 $n\times n$ 的棋盘具有 n^2 个格子。如下图所示，在所给位置里的皇后控制着同行、同列以及包含这个格子的两条斜线上的所有格子。与此对应的简单图具有 n^2 个顶点，每个顶点表示一个格子，而且若一个顶点所表示的格子里的皇后控制着另外一个顶点所表示的格子，则这两个顶点是相邻的。

皇后所控制的格子

17. 构造表示 $n\times n$ 棋盘的简单图，用边表示皇后对格子的控制，其中

    a) $n=3$　　b) $n=4$

18. 解释一下最小支配集的概念如何应用到确定控制 $n\times n$ 棋盘的最少皇后数的问题。

**19. 求控制 $n\times n$ 棋盘的最少皇后数，其中

    a) $n=3$　　b) $n=4$　　c) $n=5$

20. 假定 G_1 和 H_1 是同构的而且 G_2 和 H_2 是同构的。证明或反驳：$G_1\cup G_2$ 和 $H_1\cup H_2$ 是同构的。

21. 证明：下列性质是同构的简单图都有或都没有的不变量。

    a) 连通性　　b) 哈密顿回路的存在性　　c) 欧拉回路的存在性　　d) 有交叉数 C

    e) 有 n 个孤立顶点　　f) 是偶图

22. 如何从 G 的邻接矩阵求 $\overline{G}$ 的邻接矩阵？其中 G 是简单图。

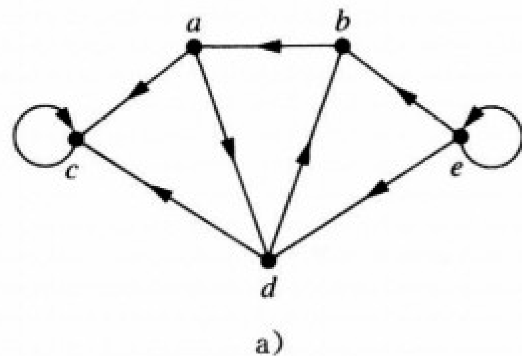
23. 有多少种非同构的、带有 4 个顶点的连通简单偶图？

*24. 有多少种非同构的简单连通图带有 5 个顶点并且

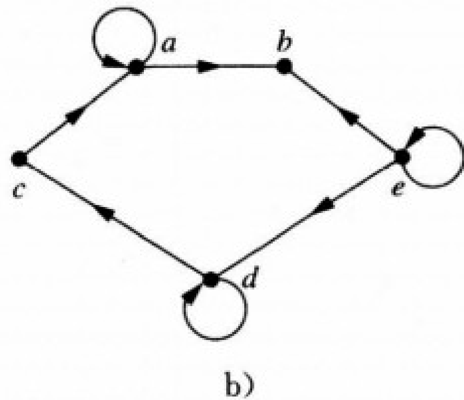
    a) 没有任何顶点的度超过 2？　　b) 色数等于 4？　　c) 非可平面的？

若有向图与它的逆同构，则它是自逆的。

25. 确定下列图是否为自逆的。



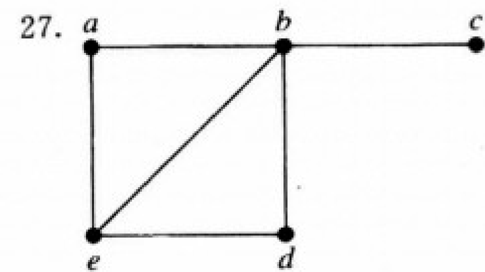
a)

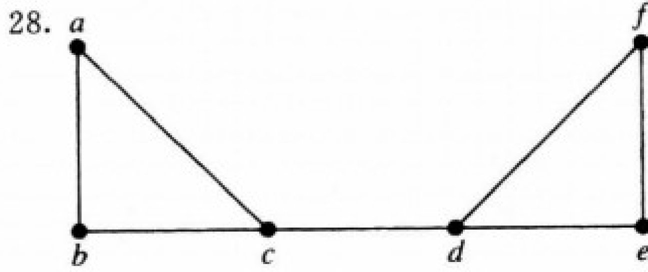


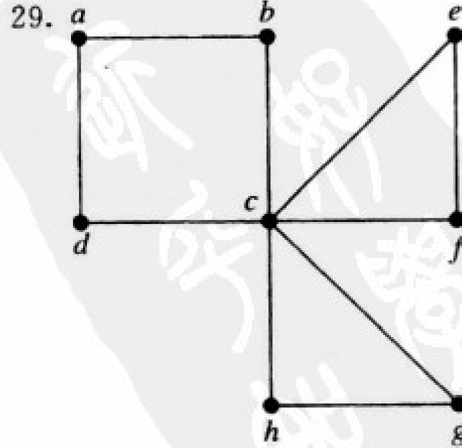
b)

26. 证明：若有向图 G 是自逆的而且 H 是同构于 G 的有向图，则 H 也是自逆的。

无向简单图的定向就是指定它的各边的方向，使得所得到的有向图是强连通的。当无向图有定向时，这个图称为可定向的。在练习 27～29 里，确定给定的图是否可定向。

27.


28.


29.


30. 因为在城市中心区交通流量正在增长，所以交通工程师正在计划把目前所有双行街道都变成单行街道。解释一下如何为这个问题建模。

*31. 证明：若一个图具有割边，则它不是可定向的。

竞赛图是简单有向图，使得若 u 和 v 是图里不同的顶点，则 (u, v) 和 (v, u) 中恰好有一个是图中的边。

32. 有多少种不同的带 n 个顶点的竞赛图？

33. 在竞赛图里一个顶点的入度与出度之和是什么？

*34. 证明：每个竞赛图都有哈密顿通路。

35. 给定鸡群里的两只鸡，其中一只占优势。这样就定义了这个鸡群的啄食次序。如何用竞赛图来为啄食次序建模？

36. 假定 G 是带 $2k$ 个奇数度顶点的连通多重图。证明：存在 k 个子图，它们的并图是 G，其中每个子图都有欧拉通路并且任何两个子图都没有公共边。(提示：向该图添加 k 条边，连接成对的奇数度顶点，并且利用这个变大了的图的欧拉回路。)

*37. 设 G 是带有 n 个顶点的简单图。G 的带宽表示成 $B(G)$，它是 $\max\{|i-j| \mid a_i$ 与 a_j 是相邻的$\}$在 G 的顶点 $a_1, a_2, \cdots, a_n$ 的所有排列上所取的最小值。即带宽是赋给相邻顶点的下标的最大差值在顶点的所有列表上所取的最小值。求下列图的带宽。

a) K_5　　b) $K_{1,3}$　　c) $K_{2,3}$　　d) $K_{3,3}$　　e) Q_3　　f) C_5

*38. 连通简单图的两个不同顶点 v_1 和 v_2 之间的距离是在 v_1 和 v_2 之间的最短通路的长度(边数)。图的半径是从顶点 v 到其他顶点的最大距离在所有顶点 v 上所取的最小值。图的直径是在两个不同顶点之间的最大距离。求下列图的半径和直径。

a) K_6　　b) $K_{4,5}$　　c) Q_3　　d) C_6

*39. a)证明：若简单图 G 的直径至少为 4，则它的补图 $\overline{G}$ 的直径不超过 2。

b)证明：若简单图 G 的直径至少为 3，则它的补图 $\overline{G}$ 的直径不超过 3。

*40. 假定一个多重图有 $2m$ 个奇数度顶点。证明：任何包含该图中每条边的回路，必然至少包含 m 条边超过一次。

41. 求 8.6 节图 8-63 中在顶点 a 与 z 之间的次最短通路。

42. 设计一个算法，它在简单连通带权图里，求在两个顶点之间的最短通路。

43. 求 8.6 节图 8-64 中在顶点 a 与 z 之间经过顶点 e 的最短通路。

44. 设计一个算法，它在简单连通带权图里，求在两个顶点之间经过第三个指定顶点的最短通路。

*45. 证明：若 G 是至少带 11 个顶点的简单图，则或者 G 不是可平面的，或者 $\overline{G}$ 不是可平面的，$\overline{G}$ 是 G 的补图。

若在图中一组顶点的集合里任何两个顶点都不相邻，则这个顶点集合称为独立的。图的独立数是图中顶点独立集里的最大顶点数。

*46. 下列图的独立数是什么？

a) K_n　　b) C_n　　c) Q_n　　d) $K_{m,n}$

47. 证明：一个简单图里的顶点数小于或等于这个图的独立数与色数之积。

48. 证明：一个图的色数小于或等于 $v-i+1$，其中 v 是这个图的顶点数，i 是这个图的独立数。

49. 假定为了生成带有 n 个顶点的随机简单图，首先选择满足 $0 \leqslant p \leqslant 1$ 的实数 p。对 $C(n, 2)$ 对不同顶点中的每一对，都生成一个在 0 与 1 之间的随机数 x。若 $0 \leqslant x \leqslant p$，则用一条边连接这两个顶点；否则就不连接这两个顶点。

a)生成带有 m 条边的图的概率是什么？其中 $0 \leqslant m \leqslant C(n, 2)$。

b)若以概率 p 来包含每一条边，则在随机生成的带有 n 个顶点的图中，边数的期望值是什么？

c)证明：若 $p=1/2$，则以相等的概率生成每一个带 n 个顶点的简单图。

每当向简单图添加更多的边(不添加顶点)时，都还保持的性质称为单调增的，每当从简单图删除边(不删除顶点)时，都还保持的性质称为单调减的。

50. 对下列每个性质来说，确定它是否为单调增的并且确定它是否为单调减的。

a)图 G 是连通的。 b)图 G 不是连通的。 c)图 G 有欧拉回路。 d)图 G 有哈密顿回路。
e)图 G 是可平面的。 f)图 G 的色数为 4。 g)图 G 的半径为 3。 h)图 G 的直径为 3。

51. 证明：图的性质 P 是单调增的当且仅当图的性质 Q 是单调减的，其中 Q 是不具有性质 P 的性质。

**52. 假定 P 是简单图的单调增的性质。证明：带 n 个顶点的随机图有性质 P 的概率是挑选一条边属于图的概率 p 的单调非减函数。

计算机题目

按下述给定的输入与输出写出程序。

1. 给定无向图的各边所关联的顶点对，确定每个顶点的度。
2. 给定有向图的各边所关联的有序顶点对，确定每个顶点的入度和出度。
3. 给定简单图的边列表，确定这个图是否为偶图。
4. 给定图的各边所关联的顶点对，构造这个图的邻接矩阵。
5. 给定图的邻接矩阵，列出这个图的各边，并且给出每条边出现的次数。
6. 给定无向图各边所关联的顶点对，以及每条边出现的次数，构造这个图的关联矩阵。
7. 给定无向图的关联矩阵，列出它的各边，并且给出每条边出现的次数。
8. 给定正整数 n，通过产生图的邻接矩阵来生成无向图，使得以相等的概率来生成所有的简单图。
9. 给定正整数 n，通过产生图的邻接矩阵来生成有向图，使得以相等的概率来生成所有的有向图。
10. 给定两个都带不超过 6 个顶点的简单图的边列表，确定这两个图是否同构。
11. 给定图的邻接矩阵和正整数 n，求顶点两两之间长度为 n 的通路数(产生对有向图和无向图来说都能工作的程序)。

*12. 给定简单图的边列表，确定它是否连通，若它不连通，则求连通分支数。

13. 给定多重图的各边所关联的顶点对，确定它是否有欧拉回路，若没有欧拉回路，则确定它是否有欧拉通路。若存在欧拉通路或欧拉回路，则构造这样的通路或回路。

*14. 给定有向多重图的各边所关联的有序顶点对，若存在欧拉通路或欧拉回路，则构造这样的通路或回路。

**15. 给定简单图的边列表，产生一条哈密顿回路，或者确定该图没有这样的回路。

**16. 给定简单图的边列表，产生一条哈密顿通路，或者确定该图没有这样的通路。

17. 给定带权连通简单图的边及其权的列表，以及该图中的两个顶点，用迪克斯特拉算法求这两点间最短通路的长度。另外，求出这条通路。
18. 给定无向图的边的表，用 8.8 节练习里所给的算法求这个图的着色。
19. 给定学生及其注册课程的表，构造期末考试日程表。
20. 给定各对电视台之间的距离，为这些台分配频率。

计算和研究

利用计算程序或你所编写的程序来做下面的练习。

1. 显示带 4 个顶点的所有简单图。
2. 显示全套的带 6 个顶点的所有非同构的简单图。
3. 显示全套的带 4 个顶点的所有非同构的有向图。
4. 随机地生成 10 个不同的简单图，每个带 20 个顶点，使得每个这样的图都是以相等的概率来生成的。
5. 构造一种格雷码，其中码字都是长度为 6 的位串。
6. 构造马在不同大小的棋盘上的周游路线。
7. 确定你在本组练习的练习 4 里生成的每个图是否为平面图。若你可以做到，则确定每个非平面图的厚度。
8. 确定你在本组练习的练习 4 里生成的每个图是否连通。若有一个图不连通，则确定这个图的连通分

支数。

9. 随机地生成带 10 个顶点的简单图。当你生成了一个带欧拉回路的图时停止。显示这个图里的一个欧拉回路。
10. 随机地生成带 10 个顶点的简单图。当你生成了一个带哈密顿回路的图时停止。显示这个图里的一个哈密顿回路。
11. 求你在本组练习的练习 4 里所生成的每个图的色数。

**12. 求旅行推销员访问美国 50 个州的每个首府所能采取的最短路线，飞机在各城市之间直线通行。

*13. 对每个不超过 10 的正整数 n 来说，估计随机生成的带 n 个顶点的简单图连通的概率，方法是生成一组随机简单图并且确定每个图是否连通。

**14. 研究这个问题：确定 $K(7, 7)$的交叉数是否为 77，79 或 81。已知它等于这三个数当中的一个。

写作题目

利用本书以外的资料，就下列问题写作短文。

1. 描述一下在 1900 年以前图论的起源和发展。
2. 讨论一下图论在生态系统研究中的应用。
3. 讨论一下图论在社会学和心理学中的应用。
4. 讨论一下通过研究网络图的性质可以了解到什么？
5. 描述一下给定一个图的顶点和边，在纸面或屏幕上画出这个图的算法。
6. 一个输入、显示和操纵各种图的软件工具应当具有什么功能？现有的工具都具有这些功能中的哪些？
7. 描述一下确定两个图是否同构的一些可用算法和这些算法的计算复杂性。目前已知最有效的算法是什么？
8. 描述一下如何用欧拉通路来帮助确定 DNA 序列。
9. 定义一下*德布鲁因序列*并且讨论它们如何出现在应用里。解释如何用欧拉回路来构造德布鲁因序列。
10. 描述一下*中国邮递员问题*并且解释如何解决这个问题。
11. 描述一下蕴含着图具有哈密顿回路的一些不同条件。
12. 描述一下用来解决旅行推销员问题的几个不同策略和算法。
13. 描述一下确定一个图是否可平面的几个不同算法。每个算法的计算复杂性是什么？
14. 在建模中，有时把大规模集成电路(VLSI)图嵌入到一本书中，让顶点都在书脊上而边都在不同的书页上(在同一页之内边与边互不相交)。定义一下图的*书页数*(book number，即上述做法所需要的最少页数——译者注)并且对 $n=3$，4，5 和 6 求包括 K_n 在内的各种图的书页数。
15. 描述一下四色定理的历史。
16. 描述一下在四色定理的证明中计算机所扮演的角色。如何肯定一个依赖计算机的证明是正确的？
17. 就是否产生最少颜色的着色以及复杂性而言，描述并比较一下给图着色的几个不同算法。
18. 解释一下在各种不同模型里如何使用图的多重着色。
19. 解释一下在带特定性质的图的非构造性存在性证明中如何使用随机图理论。

第9章　树

不包含简单回路的连通图称为树。早在 1857 年英国数学家亚瑟·凯莱就用树去计数某些类型的化合物。本章中的例子将说明从那时起，树已经被用来解决各种学科分支里的问题。

树在计算机科学里特别有用，尤其是在算法中。例如，用树来构造求出项在表中位置的有效算法；可以用树构造有效编码以节省数据存储和传输成本，比如赫夫曼编码；可以用树来研究诸如跳棋和象棋这样的博弈，并且可以帮助确定进行这些博弈的取胜策略；可以用树来为通过一系列决策而完成的过程建立模型。构造这些模型可以帮助确定比如排序算法这样的基于一系列决策的算法的计算复杂性。

通过深度优先搜索或宽度优先搜索，可以系统地遍历图的顶点，构造出一棵包括每个顶点的树来。通过深度优先搜索来探索图的顶点，也称为回溯，允许系统地搜索五花八门的问题的解，比如确定在棋盘上如何放置 8 个王后使得这些王后不能互相攻击。

可以给树的边赋权来为许多问题建立模型。例如，用加权树可以开发出构造网络的算法，这些网络含有最便宜的连接不同网络节点的电话线集合。

9.1　概述

第 8 章说明了如何用图来建立模型和解决许多问题。本章将集中讨论称为树的一种特殊类型的图，之所以这样命名是因为这样的图就像是树。例如，家族树是表示族谱图的图。家族树用顶点表示家族成员并且用边表示亲子关系。图 9-1 显示瑞士数学世家伯努利家族的家族树。表示家族树的无向图是树的一个例子。

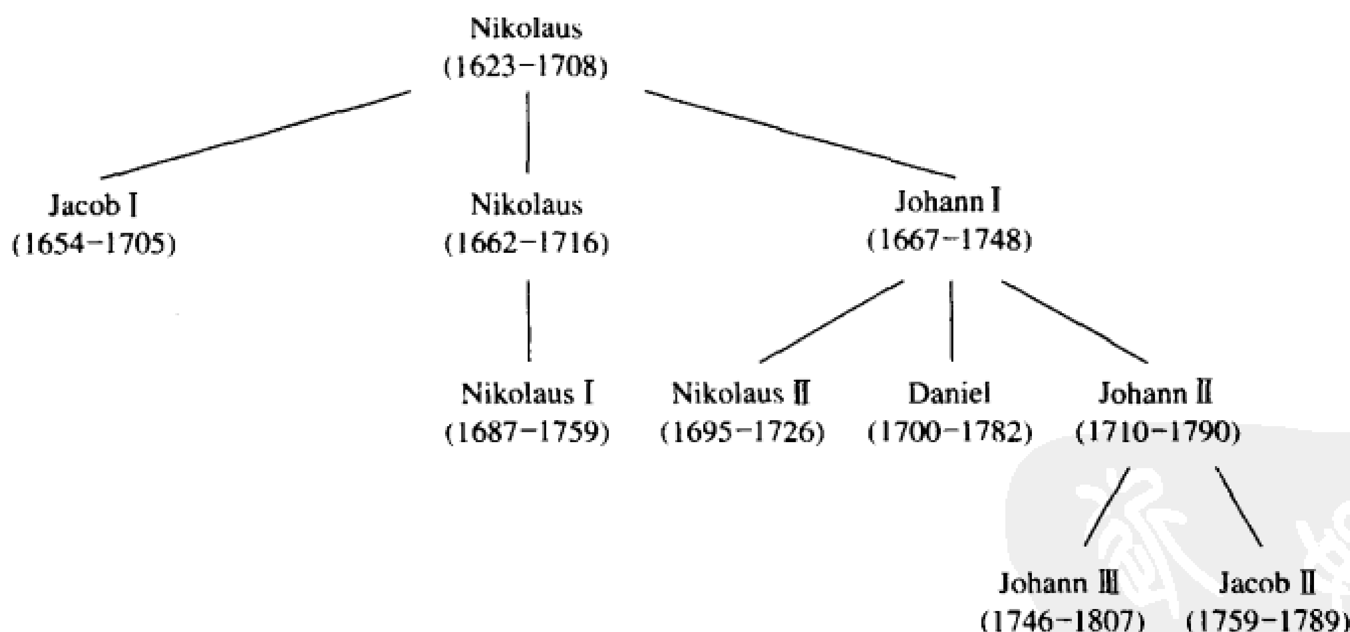


图 9-1　伯努利数学世家

定义 1　树是没有简单回路的连通无向图。

因为树没有简单回路，所以树不含多重边或环。因此任何树都必然是简单图。

例 1　在图 9-2 所示的图中，哪些图是树？

解　G_1 和 G_2 是树，因为都是没有简单回路的连通图。G_3 不是树，因为 e，b，a，d，e 是这个图中的简单回路。最后，G_4 不是树，因为它不连通。■

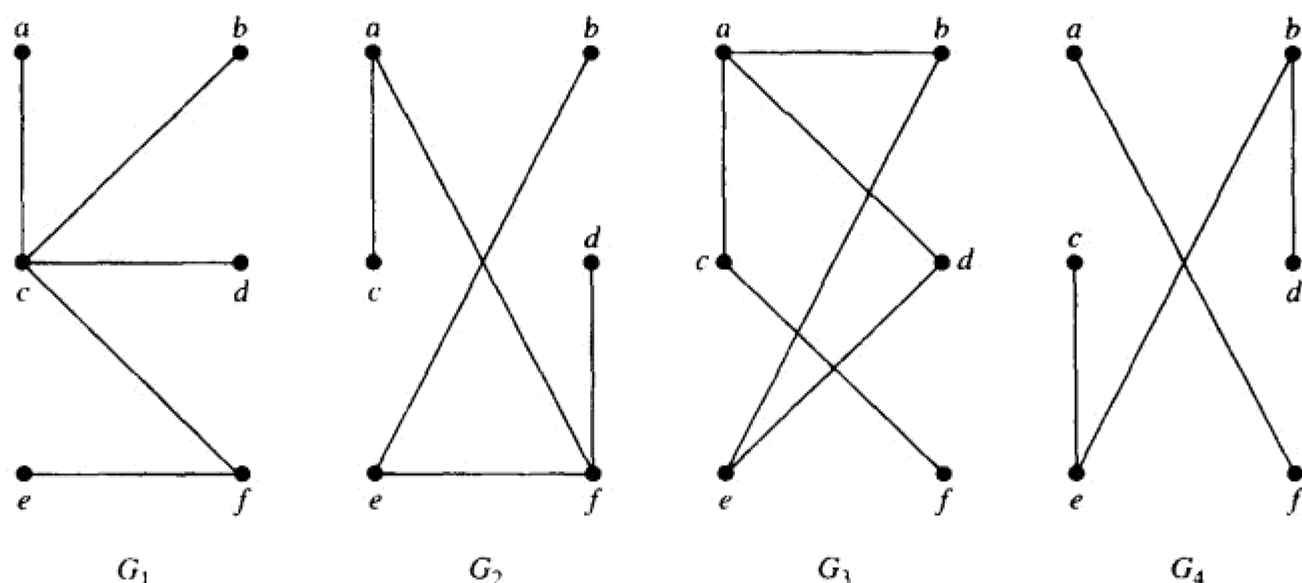


图 9-2 G_1 和 G_2 是树，G_3 和 G_4 不是树

任何一个不包含简单回路的连通图都是树。不含简单回路但不一定连通的图是什么？这些图称为**森林**，而且具有这样的性质：它们的每个连通分支都是树。图 9-3 显示一个森林。

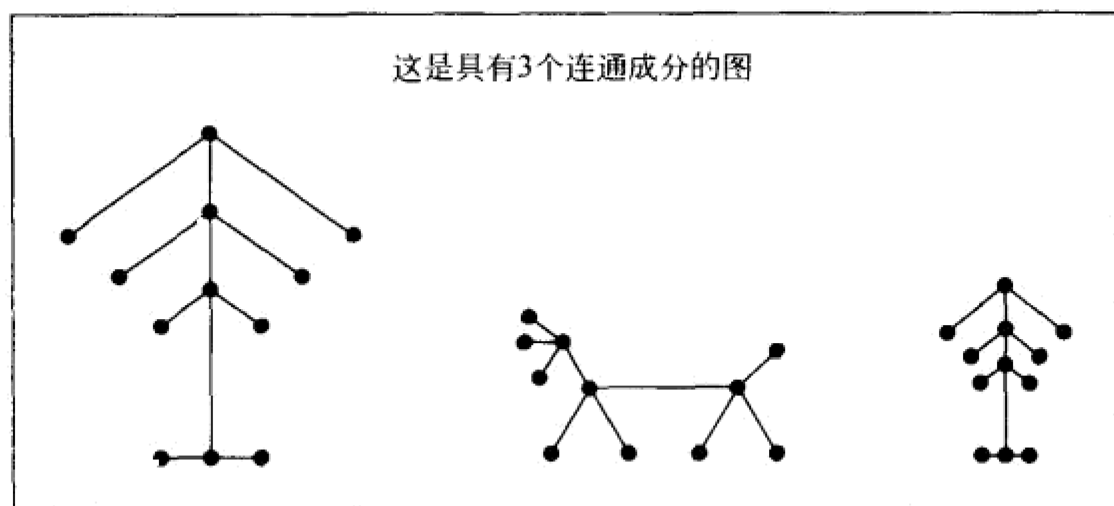


图 9-3 一个森林的例子

通常把树定义成具有在每对顶点之间存在唯一简单通路的性质的无向图。下面的定理说明这个定义与原来的定义等价。

定理 1 一个无向图是树当且仅当在它的每对顶点之间存在唯一简单通路。

证 首先假定 T 是树。则 T 是没有简单回路的连通图。设 x 和 y 是 T 的两个顶点。因为 T 是连通的，根据 8.4 节定理 1，在 x 和 y 之间存在一条简单通路。另外，这条通路必然是唯一的，因为假如存在第二条这样的通路，那么从 x 到 y 的第一条通路以及将第二条通路逆转后所得到的从 y 到 x 的通路，将组合起来形成回路。利用 8.4 节练习 43，这蕴含着在 T 中存在简单回路。因此，在树的任何两个顶点之间存在唯一简单通路。

现在假定在图 T 的任何两个顶点之间存在唯一简单通路。则 T 是连通的，因为在它的任何两个顶点之间存在通路。另外，T 没有简单回路。为了看出这是个真命题，假定 T 有包含顶点 x 和 y 的简单回路。则在 x 和 y 之间就有两条简单通路，因为这条简单回路包含一条从 x 到 y 的简单通路和一条从 y 到 x 的简单通路。因此，在任何两个顶点之间存在唯一简单通路的图是树。 □

在树的许多应用里，指定树的一个特殊顶点作为**根**。一旦规定了根，就可以给每条边指定方向如下。因为从根到图的每个顶点存在唯一通路（根据定理 1），所以指定每条边是离开根的方向。因此，树与它的根一起产生一个有向图，称为**根树**。

定义 2　根树是指定一个顶点作为根并且每条边的方向都离开根的树。

也可以递归地定义根树。参考 3.4 节来了解如何这样做。通过选择任何一个顶点来作为根，就可以把非根树变成根树。注意对根的不同选择会导致产生不同的根树。例如，图 9-4 显示通过在树 T 里分别指定 a 和 c 作为根所形成的根树。通常在画根树时把根画在图的顶端。指示根树中边的方向的箭头可以省略，因为对根的选择确定了边的方向。

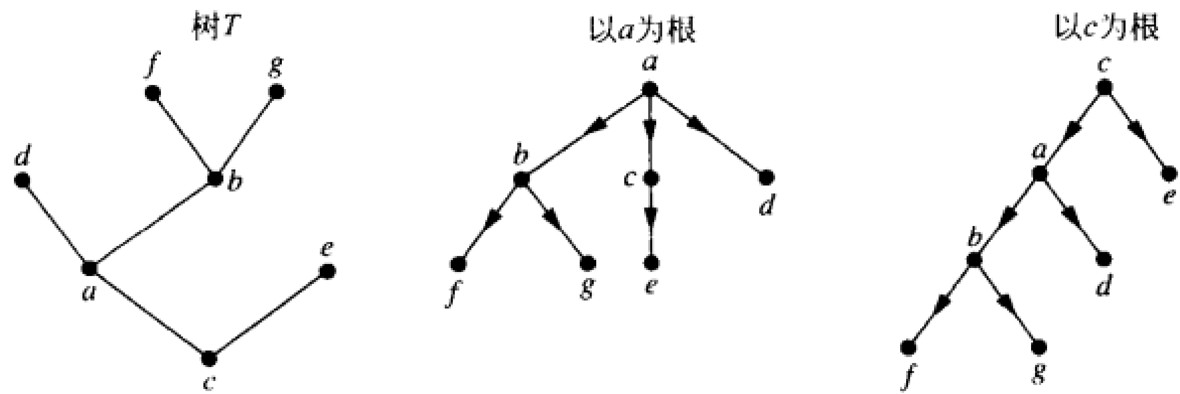


图 9-4　一棵树以及指定两个根所形成的根树

树这个术语起源于植物学和族谱学。假定 T 是根树。若 v 是 T 里非根的顶点，则 v 的父母是使得从 u 到 v 存在有向边的唯一的顶点 u（读者应当证明这样的顶点 u 是唯一的）。当 u 是 v 的父母时，v 称为 u 的子女。具有相同父母的顶点称为*兄弟*。非根顶点的*祖先*是从根到该顶点的通路上的顶点，不包括该顶点自身但包括根（即该顶点的父母，该顶点的父母的父母，等等，一直到根为止）。顶点 v 的*后代*是以 v 作为祖先的顶点。树的顶点若没有子女则称为*树叶*。有子女的顶点称为*内点*。根是内点，除非它是图中唯一的顶点，在这种情况下它是树叶。

若 a 是树里的顶点，则以 a 为根的*子树*是由 a 和 a 的后代以及这些顶点所关联的边所组成的该树的子图。

例 2　在图 9-5 所示的根树里（有根 a），求 c 的父母，g 的子女，h 的兄弟，e 的所有祖先，b 的所有后代，所有内点以及所有树叶。什么是根在 g 处的子树？

解　c 的父母是 b。g 的子女是 h，i 和 j。h 的兄弟是 i 和 j。e 的祖先是 c，b 和 a。b 的后代是 c，d 和 e。内点是 a，b，c，g，h 和 j。树叶是 d，e，f，i，k，l 和 m。根在 g 处的子树如图 9-6 所示。■

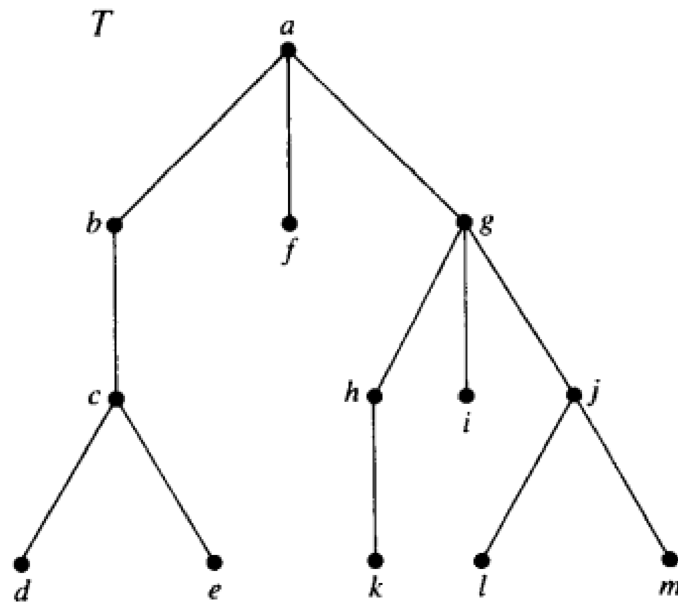


图 9-5　根树 T

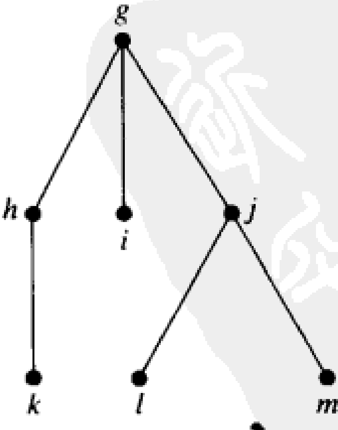


图 9-6　根在 g 处的子树

在许多不同的应用里都使用具有下面性质的根树：它们的所有内点都有同样个数的子女。在本章后面将用这样的树去研究涉及到搜索、排序和编码的问题。

定义 3 若根树的每个内点都有不超过 m 个子女，则称它为 *m 元树*。若该树的每个内点都恰好有 m 个子女，则称它为*正则 m 元树*。把 $m=2$ 的正则 m 元树称为*二叉树*。

例 3 在图 9-7 里的根树，对某个正整数 m 来说是否为正则 m 元树？

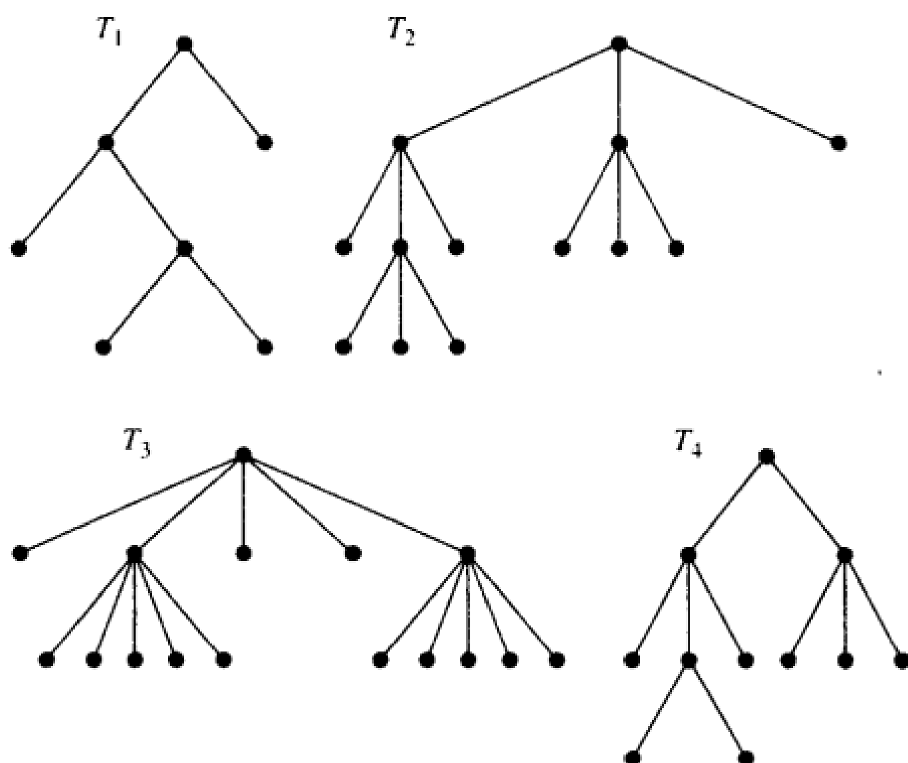


图 9-7 四个根树

解 T_1 是正则二叉树，因为它的每个内点都有两个子女。T_2 是正则三元树，因为它的每个内点都有三个子女。在 T_3 里每个内点都有五个子女，所以它是正则五元树。对任何 m 来说，T_4 都不是正则 m 元树，因为它的某些内点有两个子女而其他内点有三个子女。 ■

*有序根树*是把每个内点的子女们都排序的根树。画有序根树时，以从左向右的顺序来显示每个内点的子女们。注意在常规方式下根树的表示将确定它的边的一种顺序。我们将在画图时使用边的这种顺序，但不明确地指出认为根树是有序的。

在有序二叉树(通常只称为*二叉树*)中，若一个内点有两个子女，则第一个子女称为*左子女*而第二个子女称为*右子女*。根处在一个顶点的左子女处的树称为该顶点的*左子树*，而根处在一个顶点的右子女处的树称为该顶点的*右子树*。读者应当注意，对某些应用来说，二叉树的每个非根顶点都指定为其父母的右子女或左子女。即使当某些顶点仅有一个子女时也这样做。具体指定方式视需要而定。

可以递归地定义有序根树。3.4 节以这种方式定义了二叉树，它是有序根树的一种类型。

例 4 在图 9-8a 所示二叉树 T 里，什么是 d 的左子女和右子女(其中顺序是画法所蕴含的)？什么是 c 的左子树和右子树？

解 d 的左子女是 f 而右子女是 g。在图 9-8b 和图 9-8c 中分别显示 c 的左子树和右子树。 ■

与图的情形恰好一样，不存在用来描述树、根树、有序根树和二叉树等的标准术语。出现这种非标准的术语是因为在计算机科学里大量地使用树，而计算机科学还是相对年轻的领域。每当碰到关于树的术语时，读者就应当仔细地核对这些术语所表示的意思。

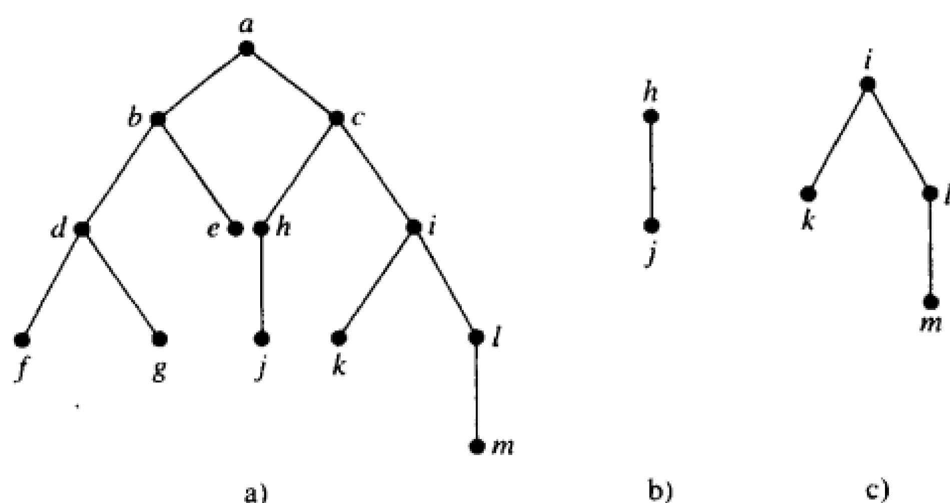


图 9-8 二叉树 T 和顶点 c 的左子树和右子树

9.1.1 树作为模型

以树为模型的应用领域非常广泛，比如计算机科学、化学、地理学、植物学和心理学等。下面将描述基于树的各式各样的模型。

例 5 饱和碳氢化合物与树 图可以用来表示分子，其中用顶点表示原子，用边表示原子之间的化学键。英国数学家亚瑟·凯莱[⊖]在 1857 年发现了树，当时他正在试图列举形如 C_nH_{2n+2} 的化合物的同分异构体，它们都称为饱和碳氢化合物。

在饱和碳氢化合物的图模型里，用 4 度顶点表示每个碳原子，用 1 度顶点表示每个氢原子。在形如 C_nH_{2n+2} 的化合物的表示图里有 $3n+2$ 个顶点。在这个图中，边数是顶点度数之和的一半。因此，在这个图中有 $(4n+2n+2)/2=3n+1$ 条边。因为这个图是连通的，而且边数比顶点数少 1，所以它必然是树(见本节末的练习 15)。

带有 n 个 4 度顶点和 $2n+2$ 个 1 度顶点的非同构的树，就表示 C_nH_{2n+2} 的不同的同分异构体。例如，当 $n=4$ 时，恰好存在两个这种类型的不同的同分异构体。它们的结构如图 9-9 所示。这两种同分异构体称为丁烷和异丁烷。 ■

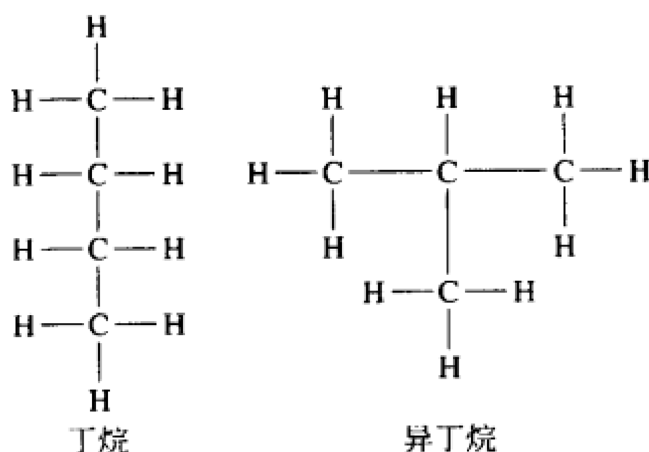


图 9-9 丁烷的两种同分异构体

例 6 表示组织机构 大的组织机构的结构可以用根树来建模。在这个树里每个顶点表示机构里的一个职务。从一个顶点到另外一个顶点的边的起点所表示的人是终点所表示的人的(直接)上司。图 9-10 所示就是这样的树。在这个树所表示的组织机构里，硬件开发主任直接为研发经理工作。 ■



⊖ **亚瑟·凯莱(Arthur Cayley, 1821—1895)** 凯莱是一个商人的儿子。他在年纪很小时就在数字计算方面显示出他的数学天才。凯莱在 17 岁时进入剑桥的三一学院。在学院期间，他爱好阅读小说。凯莱在剑桥表现优秀，被选举为任期三年的三一院士和助教。在这期间，凯莱开始了对 n 维几何学的研究，对几何学和分析学做出了多种贡献。他还喜欢登山，在瑞士度假时以登山为乐。因为没有合适的数学家职位，所以凯莱离开了剑桥大学，进入法律行业，并且在 1849 年获取律师资格。尽管凯莱在从事法律工作的同时还继续着数学研究，他依然在法律界赢得了广泛的声誉。他在法律生涯中写出了三百多篇数学论文。剑桥大学在 1863 年设立一个新的数学职位并且把它给了凯莱。凯莱接受了这份报酬低于律师的工作。

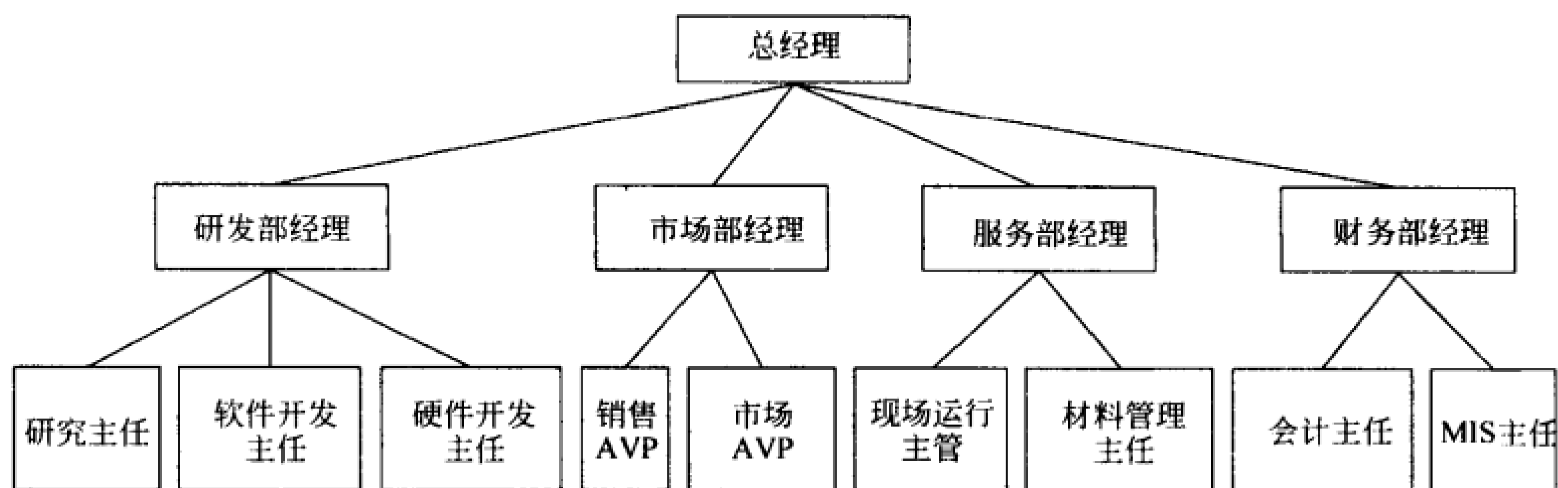


图 9-10 一家计算机公司的组织机构图

例 7 计算机文件系统 计算机存储器中的文件可以组织成目录。目录可以包含文件和子目录。根目录包含整个文件系统。因此，文件系统可以表示成根树，其中根表示根目录，内点表示子目录，树叶表示文件或空目录。在图 9-11 中显示一个这样的文件系统。在该系统中，文件 khr 属于目录 rje。(注意文件的链接，同一个文件会有一个以上的路径名，导致计算机文件系统中有回路。) ■

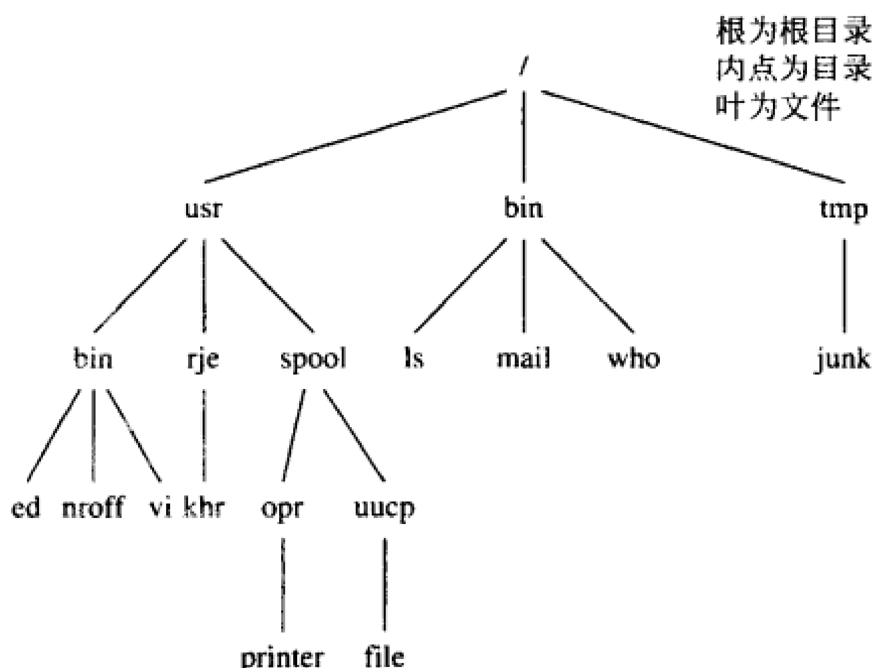


图 9-11 一个计算机文件系统

例 8 树形连接并行处理系统 在 8.2 节例 13 里描述过多种并行处理的互联网络。树形连接网络是把处理器互相连接的另外一种重要方式。这样的网络的表示图是正则二叉树。这样的网络把 $n=2^k-1$ 个处理器互连起来，其中 k 是正整数。一个非根也非树叶的顶点 v 所表示的处理器具有三个双向连接，一个连接通向 v 的父母所表示的处理器，两个连接通向 v 的两个子女所表示的处理器。根所表示的处理器具有两个双向连接，分别通向 v 的两个子女所表示的处理器。树叶所表示的处理器具有一个双向连接，通向 v 的父母。在图 9-12 里显示带 7 个处理器的树形连接网络。

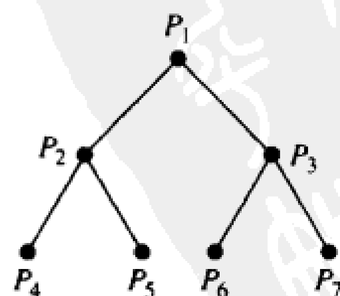


图 9-12 带 7 个处理器的树形连接网络

将要说明并行计算是如何使用树形连接网络的。具体地说，将说明图 9-12 里的处理器是如何用三步来完成 8 个数的相加的。在第一步用 P_4 相加 x_1 和 x_2。用 P_5 相加 x_3 和 x_4。用 P_6 相加 x_5 和 x_6。用 P_7 相加 x_7 和 x_8。在第二步用 P_2 相加 x_1+x_2 和 x_3+x_4。用 P_3 相加 x_5+

x_6 和 x_7+x_8。最后，在第三步用 P_1 相加 $x_1+x_2+x_3+x_4$ 和 $x_5+x_6+x_7+x_8$。这种方法要优于串行地相加 8 个数所需要的七步，串行的步骤是依次把一个数与表中前面各数之和相加。 ■

9.1.2 树的性质

我们常常需要知道树里各种各样的边和顶点的数目之间的联系。

定理 2 带有 n 个顶点的树含有 $n-1$ 条边。

证 将用数学归纳法来证明这个定理。注意对于所有的树来说，在这里可以选择一个树根并且认为这个树有根。

基础步骤：当 $n=1$ 时，有 $n=1$ 个顶点的树没有边。所有对于 $n=1$ 来说定理为真。

归纳步骤：归纳假设说有 k 个顶点的每个树都有 $k-1$ 条边，其中 k 是正整数。假设树 T 有 $k+1$ 个顶点并且 v 是 T 的树叶（v 必定存在，因为树是有穷的），设 w 是 v 的父母，从 T 删除顶点 v 和连接 w 和 v 的边，就产生有 k 个顶点的树 T'，因为所得出的图还是连通的并且没有简单回路。根据归纳假设，T' 有 $k-1$ 条边。所以 T 有 k 条边，因为 T 比 T' 多 1 条边，即连接 v 和 w 的边。这样就完成了归纳步骤。 □

下面的定理说明，带有指定的内点数的 m 元树的顶点数是确定的。与定理 2 一样，将用 n 来表示树中的顶点数。

定理 3 带有 i 个内点的正则 m 元树含有 $n=mi+1$ 个顶点。

证 除根之外的每个顶点都是内点的子女。因为每个内点有 m 个子女，所以在树中除根之外还有 mi 个顶点。因此，该树含有 $n=mi+1$ 个顶点。 □

假定 T 是完全 m 元树。设 i 是该树的内点数而 l 是树叶数。一旦 n，i 和 l 之中的一个已知，另外的两个量就随之确定了。下面的定理给出如何从已知的一个量求其他两个量的方法。

定理 4 一个正则 m 元树若带有

(i) n 个顶点，则带有 $i=(n-1)/m$ 个内点和 $l=[(m-1)n+1]/m$ 个树叶；

(ii) i 个内点，则带有 $n=mi+1$ 个顶点和 $l=(m-1)i+1$ 个树叶；

(iii) l 个树叶，则带有 $n=(ml-1)/(m-1)$ 个顶点和 $i=(l-1)/(m-1)$ 个内点。

证 设 n 表示顶点数，i 表示内点数，l 表示树叶数。利用定理 3 中的等式，即 $n=mi+1$，以及等式 $n=l+i$（这个等式为真，是因为每一个顶点要么是树叶、要么是内点），就可以证明本定理的所有三个部分。将在这里证明(i)。(ii)和(iii)的证明留给读者作为练习。

在 $n=mi+1$ 里求解 i 得出 $i=(n-1)/m$。然后把 i 的这个表达式代入等式 $n=l+i$，就证明 $l=n-i=n-(n-1)/m=[(m-1)n+1]/m$。 □

例 9 说明如何使用定理 4。

例 9 假定某人寄出一封连环信。要求收到信的每个人再把它寄给另外 4 个人。有一些人这样做了，但是其他人则没有寄出信。若没有人收到超过一封的信，而且若读过信但是不寄出它的人数超过 100 个后，连环信就终止了，则包括第一个人在内，有多少人看过信？有多少人寄出过信？

解 可以用 4 元树表示连环信。内点对应于寄出信的人，而树叶对应于不寄出信的人。因为有 100 个人不寄出信，所以在这个根树里，树叶数是 $l=100$。因此，定理 4 的(iii)说明，已经看过信的人数是 $n=(4\cdot 100-1)/(4-1)=133$。另外，内点数是 $133-100=33$，所以 33 个人寄出过信。 ■

经常需要使用这样的根树，它们是“平衡的”，所以在每个顶点处的子树都包含大约相同长度的通路。下面的一些定义将解释清楚这个概念。在根树里顶点 v 的*层*是从根到这个顶点的唯一通路的长度。根的层定义为 0。根树的*高度*就是顶点层数的最大值。换句话说，根树的层数是从根到任意顶点的最长通路的长度。

例 10 求图 9-13 所示的根树里每个顶点的层数。这棵树的高度是什么?

解 根 a 在 0 层上。顶点 b, j 和 k 都在 1 层上。顶点 c, e, f 和 l 都在 2 层上。顶点 d, g, i, m 和 n 都在 3 层上。最后，顶点 h 在 4 层上。因为任意顶点的最大层数是 4，所以这棵树的高度为 4。■

若一棵高度为 h 的 m 元树的所有顶点都在 h 层或 $h-1$ 层，则这棵树是平衡的。

例 11 在图 9-14 所示的一些根树里，哪些根树是平衡的?

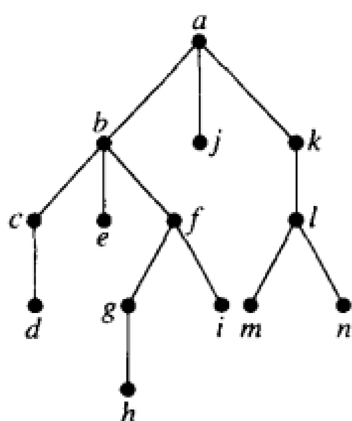


图 9-13 一棵根树

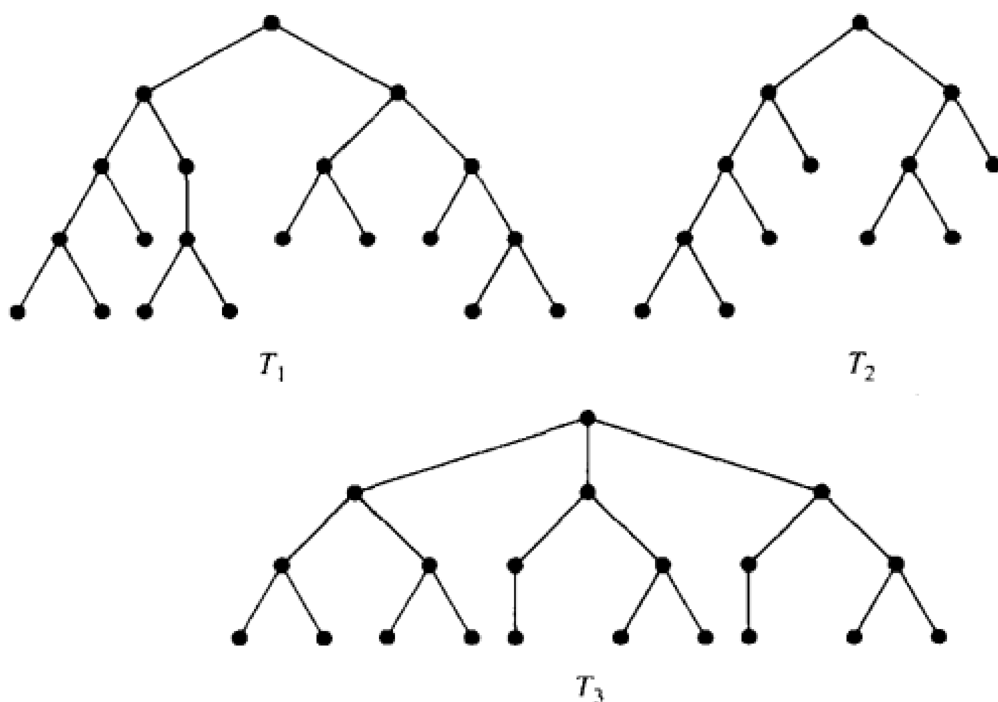


图 9-14 一些根树

解 T_1 是平衡的，因为它所有的树叶都在 3 层和 4 层上。不过，T_2 不是平衡的，因为它有树叶在 2 层、3 层和 4 层上。最后，T_3 是平衡的，因为它所有的树叶都在 3 层上。■

定理 5 中的结果建立 m 元树的高度与顶点数之间的联系。

定理 5 在高度为 h 的 m 元树里至多有 m^h 个树叶。

证 本证明对高度使用数学归纳法。首先，考虑高度为 1 的 m 元树。这些树都是由带有不超过 m 个子女的一个根所组成的，每个子女都是树叶。因此在高度为 1 的 m 元树里有不超过 $m^1=m$ 个树叶。这是归纳论证的基础步骤。

现在假定对高度小于 m 的所有 m 元树来说，这个结果都为真；这是归纳假设。设 T 是高度为 h 的 m 元树。T 的树叶都是通过删除从根到每个在 1 层的顶点的边所获得的 T 的子树的树叶，如图 9-15 所示。

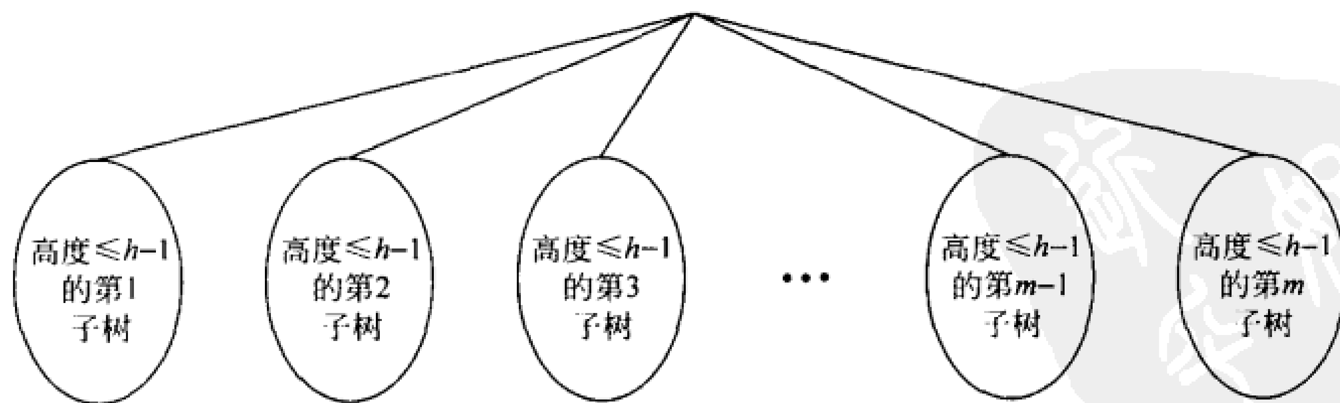


图 9-15 证明的归纳步骤

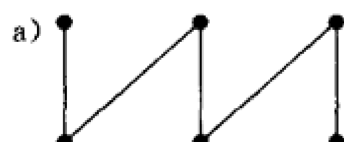
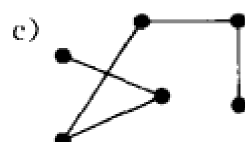
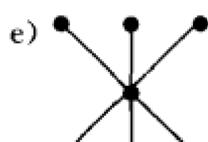
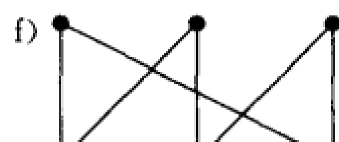
这些子树的高度都小于或等于 $h-1$。所以根据归纳假设，这些根树每个都有至多 m^{h-1} 条边。因为最多有 m 个这样的子树，每个子树最多有 m^{h-1} 个树叶，所以在这个根树里最多有 $m \cdot m^{h-1}=m^h$ 个树叶。这样就完成了归纳论证。□

推论1 若一个高度为h的m元树带有l个树叶，则$h \geqslant \lceil \log_m l \rceil$。若这个$m$元树是正则的和平衡的，则$h = \lceil \log_m l \rceil$(在这里使用向上取整函数。回忆一下，$\lceil x \rceil$是大于或等于$x$的最小整数)。

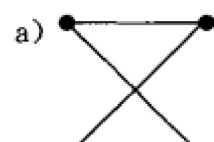
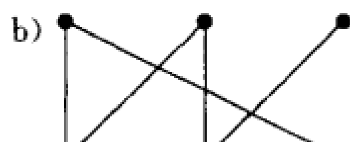
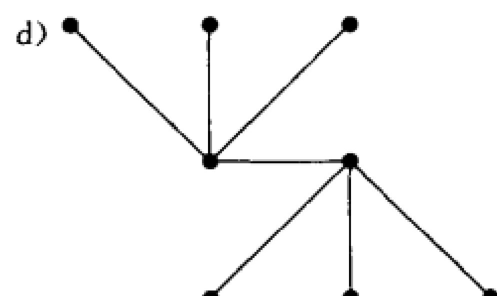
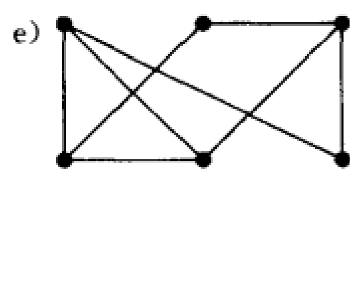
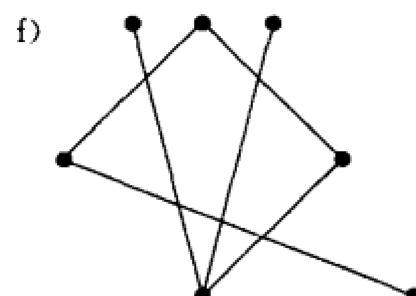
证 从定理5知道$l \leqslant m^h$。取以m为底的对数就证明$h \geqslant \log_m l$。因为h是整数，所以有$h \geqslant \lceil \log_m l \rceil$。现在假定这个树是平衡的。于是每个树叶都在$h$层或$h-1$层上，而且因为树的高度为$h$，所以在$h$层至少有一个树叶。所以必然有超过$m^{h-1}$个树叶(见本节末的练习30)。因为$l \leqslant m^h$，所以有$m^{h-1} < l \leqslant m^h$。在这个不等式里取以$m$为底的对数就得出$h-1 < \log_m l \leqslant h$。因此$h = \lceil \log_m l \rceil$。 □

练习

1. 下面哪些图是树？

a) 
b) 
c) 
d) 
e) 
f) 

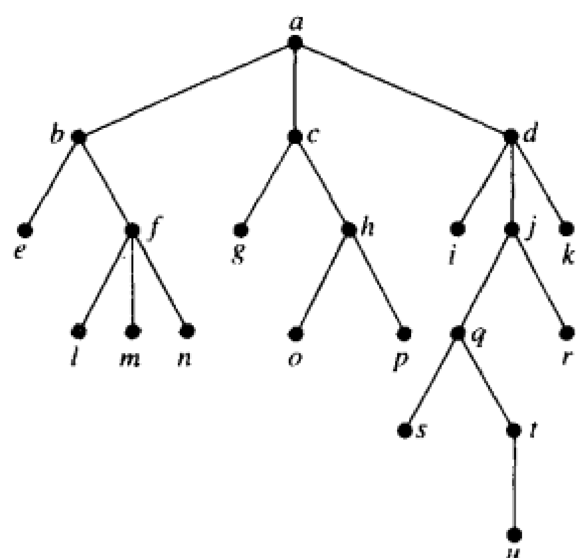
2. 下面哪些图是树？

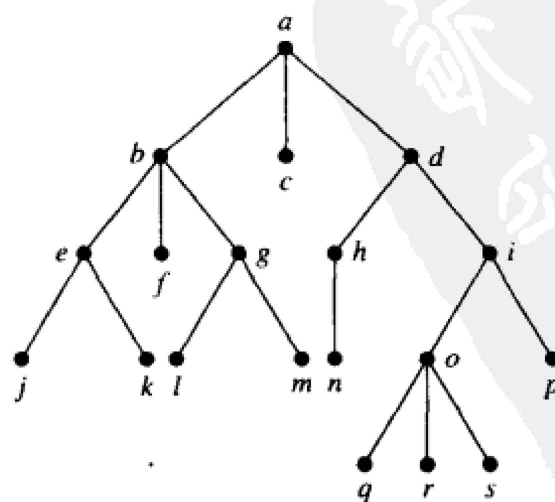
a) 
b) 
c) 
d) 
e) 
f) 

3. 回答下列关于左下图所示的根树的问题。

a)哪个顶点是根？ b)哪些顶点是内点？
c)哪些顶点是树叶？ d)哪些顶点是j的子女？
e)哪些顶点是h的父母？ f)哪些顶点是o的兄弟？
g)哪些顶点是m的祖先？ h)哪些顶点是b的后代？

4. 对于右下图所示的根树，回答练习3所列出的相同问题。





5. 练习 3 里的根树对某个正整数 m 来说，是否是正则 m 元树？

6. 练习 4 里的根树对某个正整数 m 来说，是否是正则 m 元树？

7. 练习 3 里的根树的每个顶点的层数是什么？

8. 练习 4 里的根树的每个顶点的层数是什么？

9. 画出练习 3 里的树以下列顶点为根的子树。

a) a　　b) c　　c) e

10. 画出练习 4 里的树以下列顶点为根的子树。

a) a　　b) c　　c) e

11. a)有多少种非同构的带有 3 个顶点的非根树？
b)有多少种非同构的带有 3 个顶点的根树(使用有向图的同构)？

*12. a)有多少种非同构的带有 4 个顶点的非根树？
b)有多少种非同构的带有 4 个顶点的根树(使用有向图的同构)？

*13. a)有多少种非同构的带有 5 个顶点的非根树？
b)有多少种非同构的带有 5 个顶点的根树(使用有向图的同构)？

*14. 证明：简单图是树当且仅当它是连通的、但是删除它的任何一条边就产生不连通的图。

*15. 设 G 是带有 n 个顶点的简单图。证明：G 是树当且仅当 G 是连通的并且有 $n-1$ 条边。

16. 哪些完全偶图 $K_{m,n}$ 是树，其中 m 和 n 都是正整数？

17. 带有 10 000 个顶点的树有多少条边？

18. 带有 100 个内点的正则 5 元树有多少个顶点？

19. 带有 1000 个内点的正则二叉树有多少条边？

20. 带有 100 个顶点的正则 3 元树有多少个树叶？

21. 假定 1000 个人参加象棋巡回赛。若一个选手输掉一盘就遭到淘汰，而且比赛进行到只有一位参加者还没有输过为止，则利用这个巡回赛的根树模型，来确定为了决出冠军必须下多少盘棋(假定没有平局)。

22. 一封连环信。开始时有一个人寄出一封信给其他 5 个人。收到此信的每个人要么寄出信给从来没有收到过它的其他 5 个人，要么不把它寄给任何人。假定在这个连环终止以前有 10 000 人寄出过此信，并且没有人收到过超过一封信。有多少人收到过信？又有多少人收到过信但是没有寄出它？

23. 一封连环信。开始时一个人寄出一封信给其他 10 个人。要求每个人寄出此信给其他 10 个人，而且每封信都包含该连环里前面 6 个人的列表。除非表中不足 6 个名字，否则每个人都寄一美元给表中第一个人，从表中删除这个人的名字，把其他 5 个人的名字向上移动一位，并且把他或她自己的名字插入到表的末尾。若没有人中断这个连环，并且每人至多收到一封信，则这个连环里的一个人最终将收到多少钱？

*24. 要么画出带有 76 个树叶而且高度为 3 的正则 m 元树，其中 m 是正整数，要么证明这样的树不存在。

*25. 要么画出带有 84 个树叶而且高度为 3 的正则 m 元树，其中 m 是正整数，要么证明这样的树不存在。

*26. 一个正则 m 元树 T 有 81 个树叶并且高度为 4。

a)给出 m 的上界和下界。　　b)若 T 也是平衡的，则 m 是什么？

一个完全 m 元树是其中每个树叶都在同一层上的正则 m 元树。

27. 构造高度为 4 的完全二叉树和高度为 3 的完全 3 元树。

28. 高度为 h 的完全 m 元树具有多少个顶点和多少个树叶？

29. 证明：a)定理 4 的部分(ii)。b)定理 4 的部分(iii)。

☞30. 证明：高度为 h 的正则 m 元树具有超过 m^{h-1} 个树叶。

31. 在包含总共 n 个顶点的 t 个树的森林里有多少条边？

32. 解释一下如何用树来表示由章、章中节、节中小节组成的书的目录表。

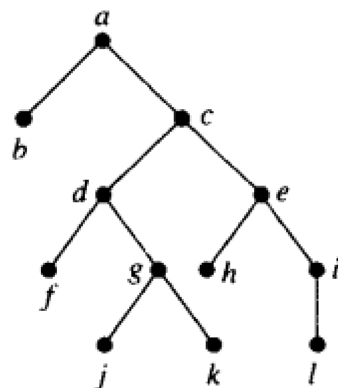
33. 下面的饱和碳氢化合物有多少种不同的同分异构体？

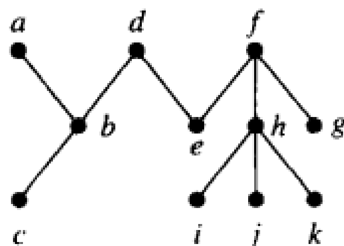
a) C_3H_8　　b) C_5H_{12}　　c) C_6H_{14}

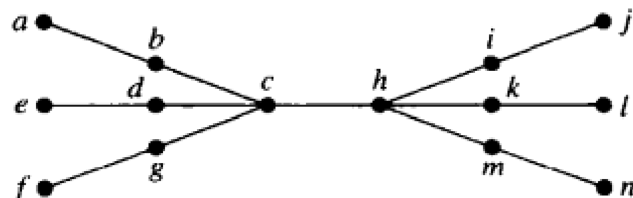
34. 在组织机构树里下述对象分别表示什么内容？
    a)一个顶点的父母。　　b)一个顶点的子女。　　c)一个顶点的兄弟。
    d)一个顶点的祖先。　　e)一个顶点的后代。　　f)一个顶点的层数。
    g)一棵树的高度
35. 对表示计算机文件系统的根树，回答与练习34里所给的那些问题相同的问题。
36. a)画出表示15个处理器的树形连接网络的带有15个顶点的完全二叉树。
    b)说明如何用(a)里的15个处理器分四步求16个数之和。
37. 设 n 是2的幂。证明：可以用 $n-1$ 个处理器的树形连接网络在 $\log n$ 步里求 n 个数之和。

*38. 标记树是其中每个顶点都指定了标记的树。当在两个标记树之间存在保持顶点标记的同构时，就把这两个标记树当作同构的。用集合{1，2，3}里三个不同的数来标记三个顶点的、非同构的标记树有多少种？用集合{1，2，3，4}里四个不同的数来标记四个顶点的、非同构的标记树有多少种？

非根树里顶点的*离心度*是从这个顶点开始的最长的简单通路的长度。若在树里没有其他顶点比一个顶点的离心度更小，则这个顶点就称为中心。在练习39～41里，求每一个是所给的树的中心的顶点。

39.


40.


41.


42. 证明：为了从非根树产生高度最小的根树，就应当选择中心来作为根。

*43. 证明：树有一个中心或两个相邻的中心。

44. 证明：每一个树都可以用两种颜色来着色。

根斐波那契树 T_n 是以下面的方式来递归地定义的。T_1 和 T_2 都是包含单个顶点的根树，而对 $n=3, 4, \cdots$ 来说，都是从一个根以 T_{n-1} 作为左子树并且以 T_{n-2} 作为右子树来构造出根树 T_n。

45. 画出前7个根斐波那契树。

*46. 根斐波那契树 T_n 有多少个顶点、树叶和内点？其中 n 是正整数。它的高度是什么？

47. 下面这个使用数学归纳法的"证明"错在什么地方？命题：有 n 个顶点的每个树都有长度为 $n-1$ 的路径。*基础步骤*：有1个顶点的每个树显然有长度为0的路径。*归纳步骤*：假设有 n 个顶点的树有长度为 $n-1$ 的路径，这个路径以 u 作为终点。加入顶点 v 和从 u 到 v 的边。所得出的树有 $n+1$ 个顶点并且有长度为 n 的路径。这样就完成了归纳步骤。

*48. 证明：有 n 个顶点的二叉树中树叶的平均深度是 $\Omega(\log n)$。

9.2 树的应用

9.2.1 引言

下面将要运用树来讨论三个问题。第一个问题是：应当如何对列表里的项进行排序，以便可以容易地找到项的位置？第二个问题是：为了在某种类型的一组对象里找出带某种性质的对象，应当做出一系列什么样的决策？第三个问题是：应当如何用位串来有效地编码一组字符？

9.2.2 二叉搜索树



在列表里搜索一些项，是计算机科学的一项重要任务。主要目标是实现一个搜索算法，当项都完全排序时，这个算法有效地找出项。这个任务可以通过使用二叉搜索树来完成，二叉搜索树是一种二叉树，其中任何顶点的每个子女都指定为右子女或左子女，没有顶点具有超过一个的右子女或左子女，而且每个顶点都用一个关键字来标记，这个关键字是各项中的一个。另外，这样指定顶点的关键字，使得顶点的关键字不仅大于它的左子树里的所有顶点的关键字，而且小于它的右子树里的所有顶点的关键字。

这个递归过程用来形成项的列表的二叉搜索树。从只包含一个顶点(即根)的树开始。指定列表中第一个项作为这个根的关键字。为了添加新的项，首先比较它与已经在树里的顶点的关键字，从根开始，若这个项小于所比较顶点的关键字而且这个顶点有左子女，则向左移动，若这个项大于所比较顶点的关键字而且这个顶点有右子女，则向右移动。当这个项小于所比较顶点的关键字而且这个顶点没有左子女时，就插入以这个项作为关键字的一个新顶点，并把新顶点作为这个顶点的左子女。同理，当这个项大于所比较顶点的关键字而且这个顶点没有右子女时，就插入以这个项作为关键字的一个新顶点，并把新顶点作为这个顶点的右子女。用例 1 来说明这个过程。

例 1 构造下面这些单词的二叉搜索树(用字母顺序)：mathematics，physics，geography，zoology，meteorology，geology，psychology 和 chemistry。

解 图 9-16 显示了构造这个二叉搜索树所用的步骤。单词 mathematics 是根的关键字。因为 physics 是在 mathematics 之后(按照字母顺序)，所以给根添加带关键字 physics 的右子女。因为 geography 是在 mathematics 之前，所以给根添加带关键字 geography 的左子女。下一步，给带关键字 physics 的顶点添加右子女，并且给其指定关键字 zoology，因为 zoology 是在 mathematics 之后且在 physics 之后。同理，给带关键字 physics 的顶点添加左子女，并且给其指定关键字 meteorology。给带关键字 geography 的顶点添加右子女，并且给其指定关键字 geology。给带关键字 zoology 的顶点添加左子女，并且给其指定关键字 psychology。给带关键字 geography 的顶点添加左子女，并且给其指定关键字 chemistry(读者应当完成在每步上所需的所有比较)。 ■

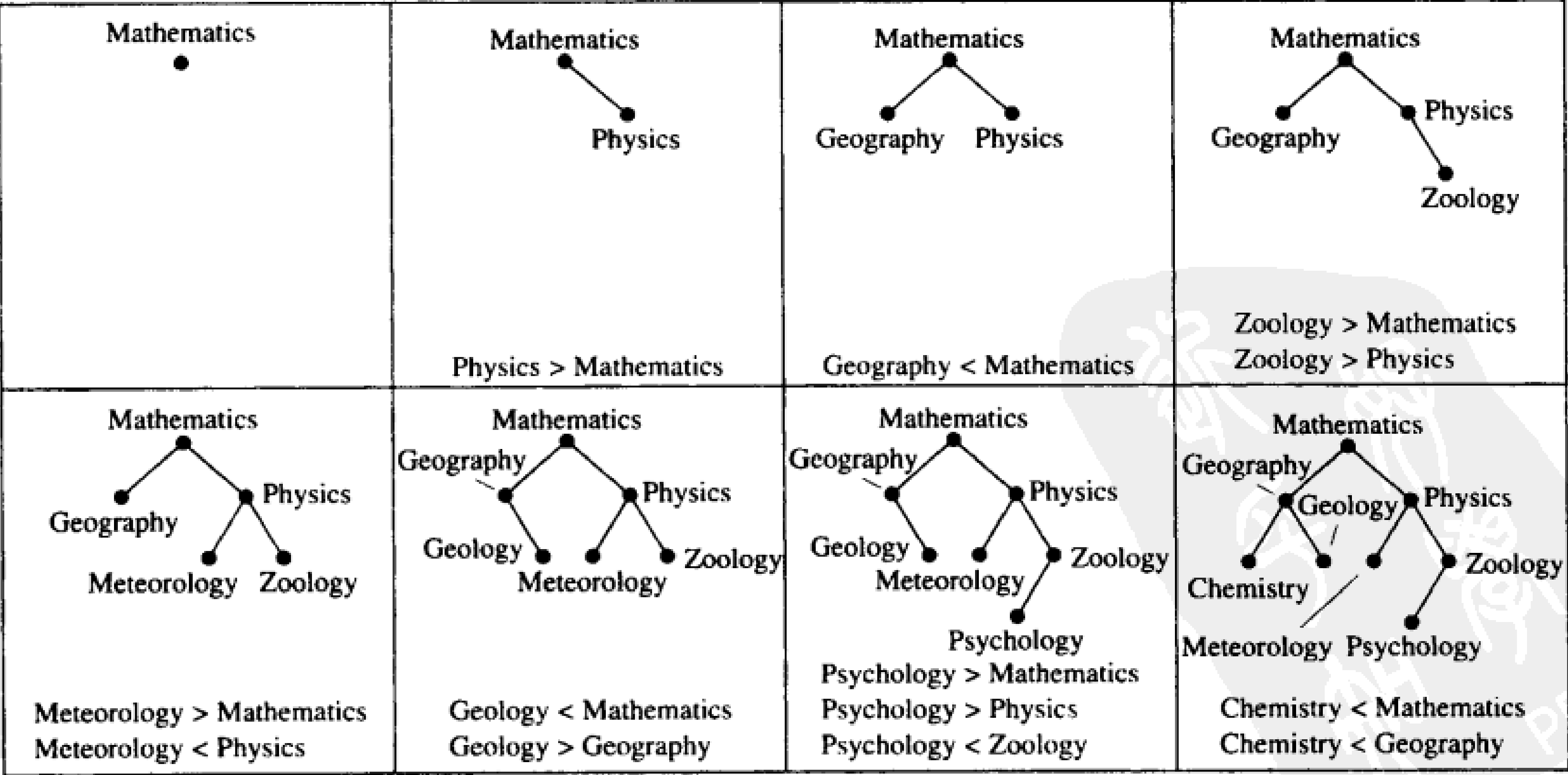


图 9-16 构造二叉搜索树

为了找出项的位置，我们尝试将它添加到二叉搜索树。若它出现，则找到它的位置。算法 1 给出这样的伪代码，它在二叉搜索树寻找一个项的位置，若没有找到这个项，则添加一个新的顶点，用这个项作为新顶点的关键字。当 x 不是关键字时，给树添加带关键字 x 的新顶点。在这个伪代码中，v 是以 x 作为其关键字的顶点，而 $label(v)$ 表示顶点 v 的关键字。

```
算法 1  二叉搜索树算法
procedure insertion(T: 二叉搜索树, x: 项)
v := T 的根
{一个不在 T 里出现、具有值 null 的顶点}
while v≠null 并且 label(v)≠x
begin
   if x<label(v) then
   if v 的左子女≠null then v := v 的左子女
   else 添加 new vertex 作为 v 的左子女并且设置 v := null
   else
   if v 的右子女≠null then v := v 的右子女
   else 给 T 添加 new vertex 作为 v 的右子女并且设置 v := null
end
if T 的根=null then 给树添加顶点 r 并且用 x 标记它
else if label(v)≠x then 用 x 标记 new vertex
{v=x 的位置}
```

现在我们来确定这个过程的计算复杂性。假定有 n 个项的列表的二叉搜索树 T。可以从 T 这样构造一个满二叉树 U：在必要时添加无标记的顶点，以使得每个带关键字的顶点都有两个子女。这个做法在图 9-17 里说明。一旦这样做了，就容易找出新项的位置，或者添加新项作为关键字而不添加顶点。

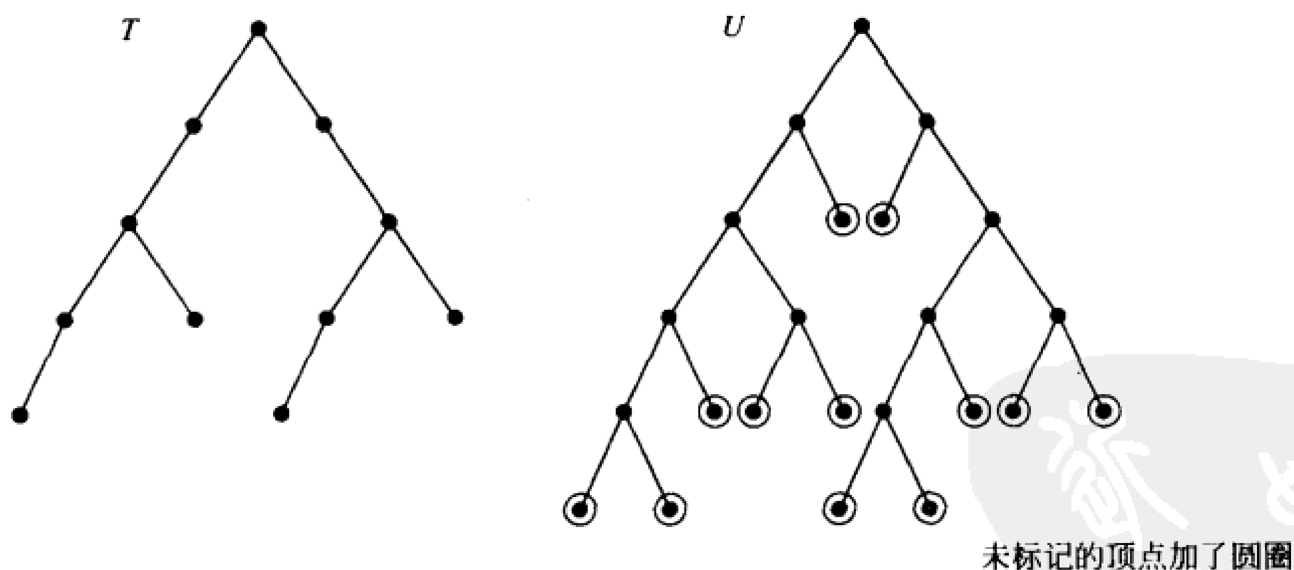


图 9-17 添加无标记顶点使二叉搜索树满

添加一个新项所需要的比较次数，最多等于在 U 里从根到树叶的最长通路的长度。U 的内点都是 T 的顶点。所以 U 有 n 个内点。现在可以利用 9.1 节定理 4 的部分(ii)来得出 U 有 $n+1$ 个树叶。利用 9.1 节的推论 1，可以看出 U 的高度大于或等于 $h=\lceil\log(n+1)\rceil$。所以，必须至少执行 $\lceil\log(n+1)\rceil$ 次比较以添加某个项。注意若 U 是平衡的，则它的高度是 $\lceil\log(n+1)\rceil$（根据

9.1 节的推论 1)。因此，若二叉搜索树是平衡的，则确定一个项的位置或者添加一个项所需要的比较次数不超过$\lceil \log(n+1) \rceil$次。当给二叉搜索树添加一些项时，该树可能变得不平衡。因为平衡的二叉搜索树给出二叉搜索的最优的最坏情形复杂性，所以添加项时重新平衡二叉搜索树的算法已经设计出来。感兴趣的读者可以查阅关于数据结构的参考文献来了解这些算法。

9.2.3 决策树

根树可以用来为这样的问题建立模型，其中一系列决策将生成一个解。例如，二叉搜索树可以用来基于一系列比较来找出项的位置，其中每次比较都说明是否已经找到了项的位置，或者是否应当向右或向左进入子树。其中每个内点都对应着一次决策，这些顶点的子树都对应着该决策的每种可能后果，这样的根树称为决策树。问题的可能的解对应着这个根树的通向树叶的通路。下一个例子说明决策树的应用。

例 2 假定有重量相同的 7 枚硬币和重量较轻的一枚伪币。为了用一架天平确定这 8 枚硬币中哪个是伪币，需要多少次称重？给出找出这个伪币的算法。

解 在天平称上每次称重结果有三种可能性。分别是：两个托盘有相同的重量，第一个托盘较重，或第二个托盘较重。所以，称重序列的决策树是 3 元树。在决策树里至少有 8 个树叶，这是因为有 8 种可能的后果(因为每枚硬币都可能是较轻的伪币)，而每种可能的后果必须至少用一个树叶来表示。确定伪币所需要的最大称重次数是决策树的高度。从 9.1 节的推论 1 得出决策树的高度至少是$\lceil \log_3 8 \rceil = 2$。因此，至少需要两次称重。

用两次称重来确定伪币，是可行的。说明如何这样做的决策树如图 9-18 所示。 ■

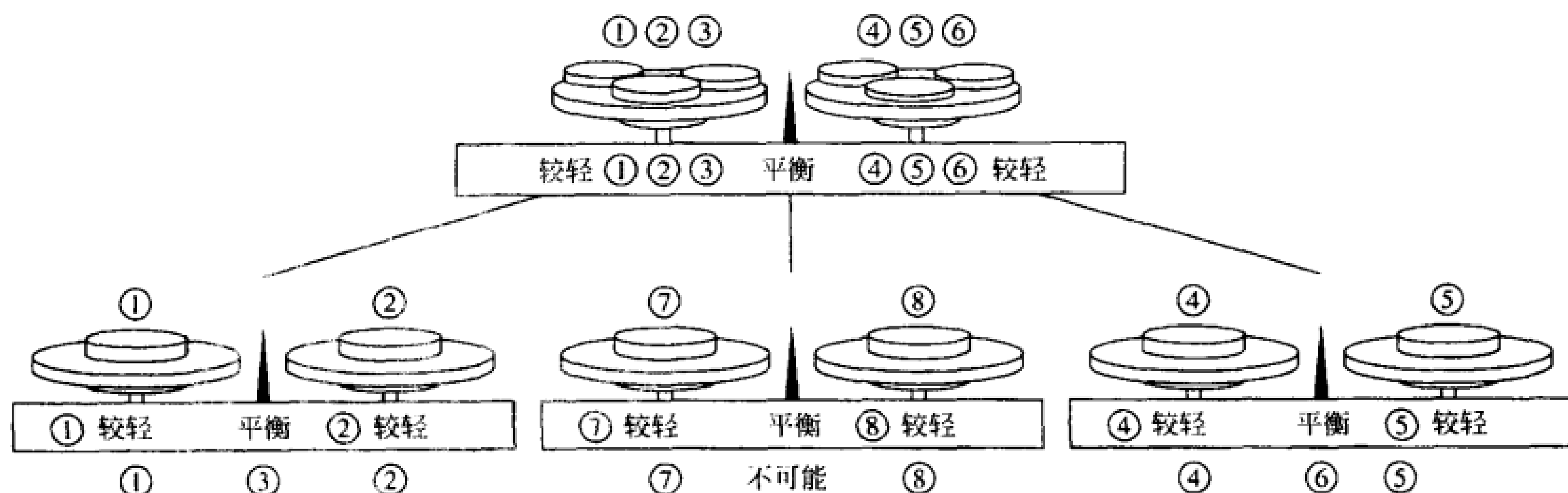


图 9-18 找出伪币位置的决策树。最下面一行为伪币判断结果

排序算法的复杂性 已经开发了许多不同的排序算法。为了确定一个具体的排序算法是否有效，就要确定这个算法的复杂性。用决策树作为模型，可以求出排序算法的最坏情形复杂性的下界。

可以用决策树为排序算法建立模型并且确定对这些算法的最坏情形复杂性的估计。注意给定 n 个元素，这些元素有 $n!$ 种可能的排序，因为这些元素 $n!$ 种排列中的每一个都可以是正确的顺序。本书研究的排序算法以及最常用的排序算法都是基于二叉比较，即一次比较两个元素。每次这样的比较都缩小了可能顺序的集合。因此，基于二叉比较的排序算法可以表示成二叉决策树，其中每个内点表示两个元素的一次比较。每个树叶表示 n 个元素的 $n!$ 种排列中的一种。

例 3 在图 9-19 里显示排序列表 a，b，c 里元素的决策树。 ■

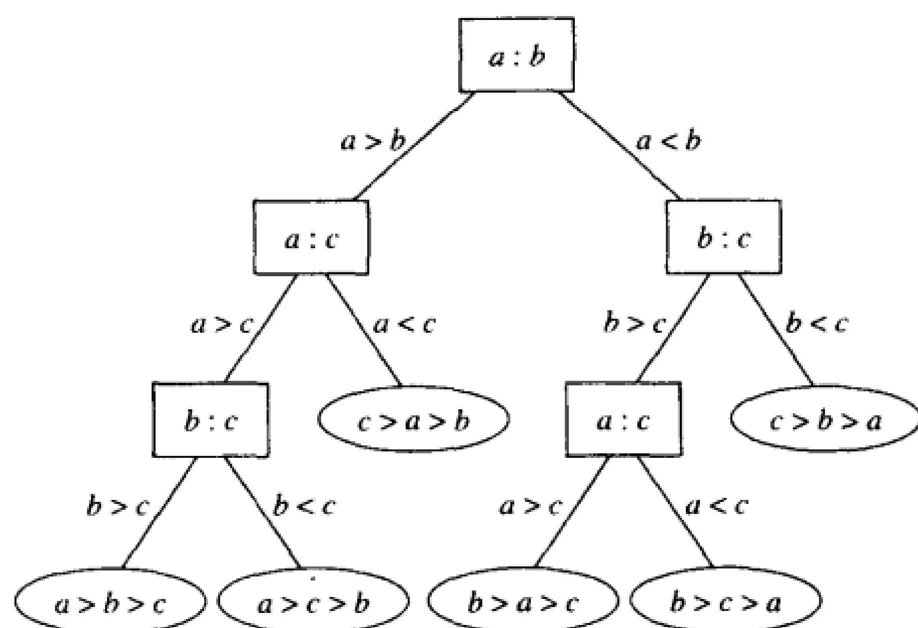


图 9-19 排序三个不同元素的决策树

基于二叉比较的排序的复杂性是用二叉比较的次数来度量的。排序有 n 个元素的列表所需要的最多比较次数就给出了这个算法的最坏情形复杂性。所用的最多比较次数等于表示这个排序过程的决策树里的最长通路长度。换句话说，所需要的最多比较次数等于这个决策树的高度。因为带 $n!$ 个树叶的二叉树的高度至少是 $\lceil \log n! \rceil$（利用 9.1 节推论 1），所以定理 1 说至少需要 $\lceil \log n! \rceil$ 次比较。

定理 1 基于二叉比较的排序算法至少需要 $\lceil \log n! \rceil$ 次比较。

可以用定理 1 来给出基于二叉比较的排序算法所用比较次数的大 O 估计。只需要注意到 $\lceil \log n! \rceil$ 是 $\Theta(n\log n)$，这是算法的计算复杂性经常使用的一个参照函数。这个估计是正确的，因为根据 2.2 节例 5 有 $\lceil \log n! \rceil$ 是 $O(n\log n)$。根据 2.2 节练习 62 有 $\lceil \log n! \rceil$ 是 $\Omega(n\log n)$，因为对于 $n>4$ 来说 $\log n! > (n\log n)/4$。推论 1 是这个估计的结果。

推论 1 基于二叉比较的排序算法排序 n 个元素所用的比较次数是 $\Omega(n\log n)$。

推论 1 的一个后果是，基于二叉比较的排序算法在最坏情形下使用 $\Theta(n\log n)$ 次比较来排序 n 个元素，其他这类算法都没有更好的最坏情形复杂性，在这个意义下，基于二叉比较的排序算法是最优的。注意根据 3.5 节定理 1 就看出，在这个意义下归并排序算法是最优的。

对于排序算法的平均情形复杂性也可以证明类似的结果。基于二叉比较的排序算法所用的平均比较次数是表示这个排序算法的决策树中的平均树叶深度。根据 9.1 节练习 48 就知道，有 N 个顶点的二叉树的平均树叶深度是 $\Omega(\log N)$。当令 $N=n!$ 并且注意到因为 $\log n! = \Theta(n\log n)$，所以是 $\Omega(\log n!)$ 的函数也是 $\Omega(n\log n)$ 时，就会得出下面的估计。

定理 2 基于二叉比较的排序算法排序 n 个元素所用的平均比较次数是 $\Omega(n\log n)$。

9.2.4 前缀码

考虑一下这样的问题：用位串来编码英语字母表里的字母（其中不区分小写和大写字母）。可以用长度为 5 的位串来表示每个字母，因为只有 26 个字母而且有 32 个长度为 5 的位串。当每个字母都用 5 位来编码时，用来编码数据的总位数是 5 乘以文本中的字符数。有没有可能找出这些字母的编码方案，使得在编码数据时使用的位更少？若可能，那么就可以节省存储空间而且缩短传输时间。

考虑用不同长度的位串来编码字母。较短的位串用来编码出现较频繁的字母，较长的位串应当用来编码不经常出现的字母。当用变化的位数来给字母编码时，就必须用某种方法来确定

每个字母的位在何处开始和结束。例如，若把 e 编码成 0，把 a 编码成 1，而把 t 编码成 01，则位串 0101 可能对应着 *eat*、*tea*、*eaea* 或 *tt*。

为了保证没有位串对应着多个字母的序列，可以令一个字母的位串永远不出现在另外一个字母的位串的开头部分。具有这个性质的编码称为*前缀码*。例如，把 e 编码成 0、把 a 编码成 10、而把 t 编码成 11 的编码就是前缀码。从编码着一个单词的字母的唯一位串可以恢复这个单词。例如，串 10110 是 *ate* 的编码。为了看明白这一点，注意开始的 1 不表示一个字符，但是 10 表示 a(并且它不可能是另外一个字母的位串的开始部分)。然后，下一个 1 不表示一个字符，但是 11 表示 t。最后一位 0 表示 e。

前缀码可以用二叉树来表示，其中字符是树里树叶的标记。标记树的边，使得通向左子女的边标记为 0 而通向右子女的边标记为 1。用来编码一个字符的位串是在从根到以这个字符作为标记的树叶的唯一通路上边的标记的序列。例如，图 9-20 里的树表示把 e 编码成 0，把 a 编码成 10，把 t 编码成 110，把 n 编码成 1110 和把 s 编码成 1111。

表示编码的树可以用来解码位串。例如，考虑一个用图 9-20 里的编码编成 11111011100 的单词。这个位串可以这样解码：从根开始，用位的序列来形成一条到树叶为止的通路。每个 0 位都使得通路向下到达通向通路里上一个顶点的左子女的边，而每个 1 位都对应到上一个顶点的右子女。所以，开头的 1111 对应这样的通路：从根开始，向右前进四次，到达以 s 作为标记的树叶，因为串 1111 是 s 的编码。从第五个位继续进行，在向右再向左之后，就到达下一个树叶，这时访问以 a 作为标记的顶点，其中把 a 编码成 10。从第七个位开始，在向右三次然后向左之后，到达了下一个树叶，这时访问用 n 标记的顶点，其中把 n 编码成 1110。最后，末位 0 通向用 e 标记的树叶。因此，原来的单词是 *sane*。

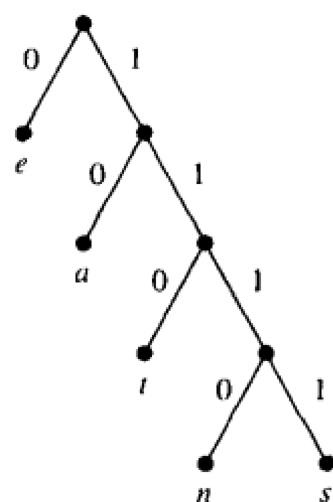


图 9-20　表示前缀码的二叉树

可以从任何二叉树来构造一个前缀码，其中每个内点的左边都用 0 标记，而右边都用 1 标记，树叶都用字符标记。字符都用从根到这个树叶的唯一通路里的边的标记所组成的位串来编码。

赫夫曼编码　现在介绍一种算法，这种算法用一个字符串中符号的出现频率(即出现概率)作为输入，并且产生编码这个字符串的一个前缀码作为输出，在这些符号的所有可能的二叉前缀码中，这个编码使用最少的位。这个所谓*赫夫曼编码*的算法是大卫·霍夫赫[1] 1951 年做麻省理工学院的研究生时发表在一篇学期论文中的。(注意这个算法假定已知字符串中每个符号出现多少次，所以可以计算每个符号的出现频率，方法是用

[1] **大卫·赫夫曼(David A. Huffman, 1925—1999)**　赫夫曼在俄亥俄州长大。他 18 岁时于俄亥俄州立大学获得电机工程的学士学位。此后他在美国海军服役，担任一艘扫雷舰上的雷达维护官，第二次世界大战后在亚洲水域执行扫雷任务。后来，他从俄亥俄州立大学获得硕士学位并且从麻省理工学院获得电机工程的博士学位。赫夫曼在 1953 年成为麻省理工学院的教员并一直呆在那里，直到 1967 年成为加州大学桑塔·克鲁茨分校计算机科学系的创办人之一。他在该系的发展中起到了重要作用并且在那里度过了余下的职业生涯。他 1994 年退休。

赫夫曼为人所知，是由于他在信息论与编码、雷达与通信的信号设计、异步逻辑电路的设计过程等方面的贡献。赫夫曼在 0 曲率表面上的工作导致他开发出把纸和乙烯基塑料折叠成非同寻常的形状的独创技术。这些形状被许多人当作艺术品并且在几个展览中公开展示。但是，赫夫曼最出名的，还是由于他开发了现在所谓的赫夫曼编码。这是他在麻省理工学院做研究生期间所写的一篇学期论文中的一个结果。

赫夫曼喜爱户外探险，经常远足和旅游。当他快 70 岁时还获得了水肺潜水员的资格。他饲养毒蛇作为宠物。

这个符号出现的次数除以这个字符串的长度。)赫夫曼编码是数据压缩中的基本算法，数据压缩的目的在于减少表示信息所需要的位数。赫夫曼编码广泛用于压缩表示文本的位串，并且在压缩视频和图像文件方面也起到重要作用。

算法 2 表示赫夫曼编码算法。给定符号及其频率，目标是构造一个根二叉树，其中符号是树叶的标记。算法从一些树的一个森林开始，每个树含有一个顶点，每个顶点有一个符号作为标记，并且这个顶点的权就等于所标记符号的频率。在每一步，都把具有最小总和权的两个树组合成一个单独的树，方法是引入一个新的根，把具有较大的权的树作为左子树，把具有较小的权的树作为右子树。另外，把这个树的两个子树的权之和作为这个树的总的权。(虽然可以规定在具有相同的权的树之间进行选择以打破平局的过程，但是这里将不具体指定这样的过程。)当构造出了一个树，即森林缩小为单个树时，算法就停止。

算法 2 赫夫曼编码

procedure $Huffman$(C: 具有频率 w_i 的符号 a_i, $i=1, 2, \cdots, n$)
$F:=n$ 个根树的森林，每个根树由单个顶点 a_i 组成并且赋权 w_i
while F 不是树
begin
  把 F 中满足 $w(T)\geq w(T')$ 的权最小的根树 T 和 T' 换成具有新树根的一个树，
    这个树根以 T 作为左子树并且以 T' 作为右子树。
    用 0 标记树根到 T 的新边，并且用 1 标记树根到 T' 的新边。
  把 $w(T)+w(T')$ 作为新树的权。
end
{符号 a_i 的赫夫曼编码是从树根到 a_i 的唯一路径上的边的标记的连接}

例 4 说明如何用算法 2 来编码一组 5 个符号。

例 4 用赫夫曼编码来编码下列符号，这些符号具有下列频率：A：0.08，B：0.10，C：0.12，D：0.15，E：0.20，F：0.35。编码一个字符串所需要的平均位数是多少？

解 图 9-21 表示了编码这些符号所用的步骤。所产生的编码为：A 是 111，B 是 110，C 是 011，D 是 010，E 是 10，F 是 00。使用这种编码来编码一个符号所用的平均位数是 $3\cdot 0.08+3\cdot 0.10+3\cdot 0.12+3\cdot 0.15+2\cdot 0.20+2\cdot 0.35=2.45$。■

注意赫夫曼编码是贪心算法。在每一步替换具有最小权的两棵树，在没有任何二叉前缀码能使用更少的位来编码这些符号的情况下，这样做就导出了最优编码。在本节末把赫夫曼编码是最优的证明留作练习 32。

赫夫曼编码有许多变种。例如，不编码单个符号，可以编码指定长度的符号块，比如两个符号的块。这样做有可能减少编码这个字符串所需要的位数(参看本节末的练习 30)。也可以用两个以上的符号来编码这个符号串中的原始符号(参看本节末练习 28 前面的说明)。另外，当事先不知道一个字符串中每个符号的频率时，可以使用一种变种，即所谓的自适应赫夫曼编码(参见[Sa00])，使得在读这个字符串的同时来进行编码。

9.2.5 博弈树



可以用树来分析某些类型的游戏，比如圈叉游戏、轮流取石头游戏、跳棋和象棋。在每一种游戏当中，两个选手轮流进行移动。每个选手知道另一个选手的移动并且游

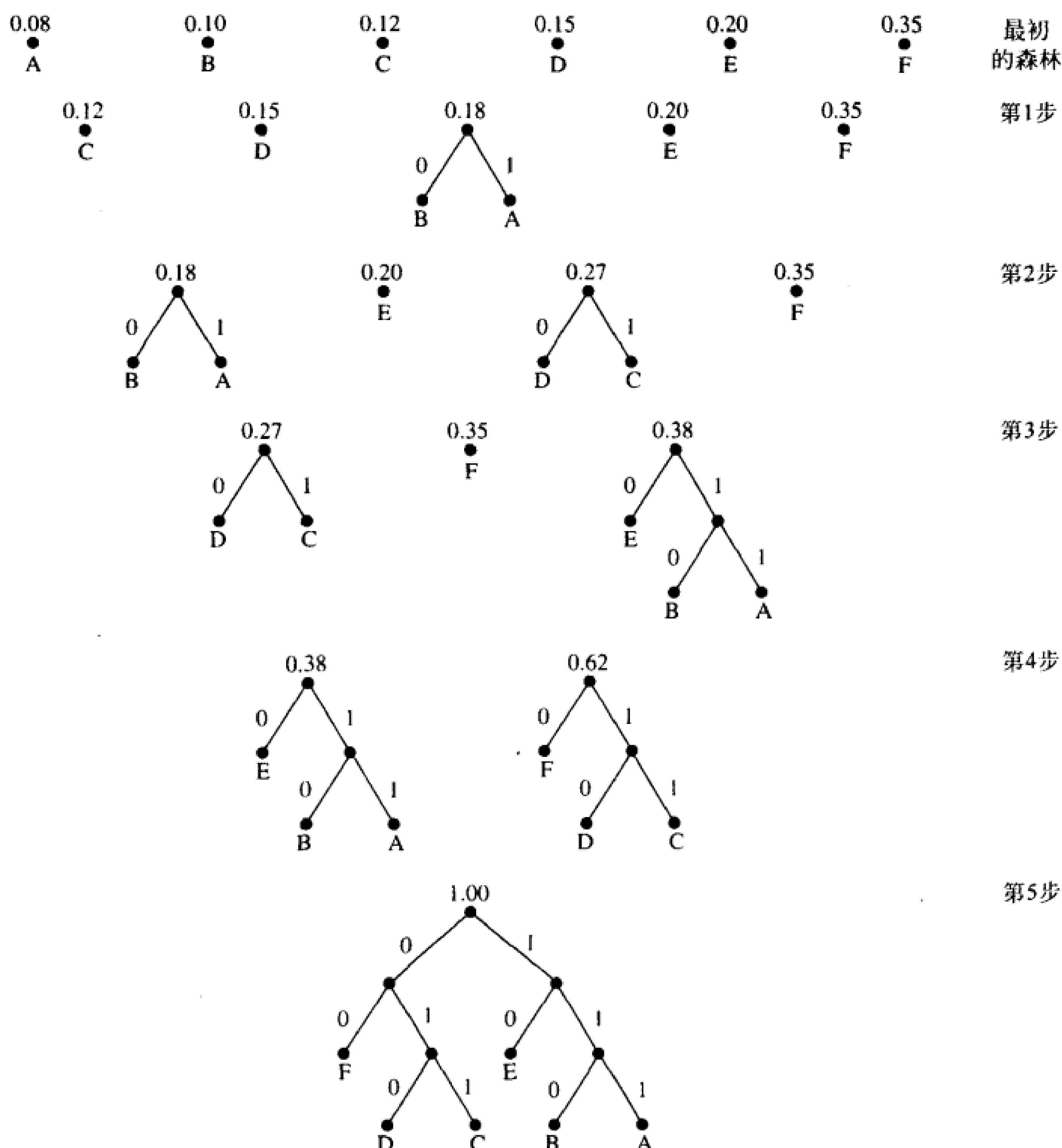


图 9-21 例 4 中符号的赫夫曼编码

戏里不存在偶然因素。使用博弈树为这样的游戏建立模型，这些树的顶点表示当游戏进行时游戏所处的局面，边表示在这些局面之间合乎规则的移动。由于博弈树常常很大，所以通过用同一个顶点表示所有对称的局面来简化博弈树。但是，如果一个游戏的不同移动序列导致同一个局面，则可以用不同的顶点来表示这个局面。通常的约定是用方框表示偶数层的顶点并且用圆圈表示奇数层的顶点。当游戏处在偶数层顶点所表示的局面时，就轮到第一个选手移动。当游戏处在奇数层顶点所表示的局面时，就轮到第二个选手移动。博弈树所表示的游戏可以永远不结束，比如进入了无穷循环，因而博弈树可以是无穷的，但是对于大多数游戏来说，都存在一些规则导致有穷的博弈树。

博弈树的树叶表示游戏的终局。给每个树叶指定一个值，来表示当游戏在这个树叶所代表的局面里终止时第一个选手的得分。对于非胜即负的游戏，用 1 来标记圆圈所表示的终结顶点以表示第一个选手获胜，用 -1 来标记方框所表示的终结顶点以表示第二个选手获胜。对于允许平局的游戏，用 0 来标记平局所对应的终结顶点。注意，对于非胜即负的游戏，为终结顶点指定

值，这个值越高，第一个选手的结局就越好。

例 5 展示一个非常著名的和经过深入研究的游戏的博弈树。

例 5 轮流取石头游戏 轮流取石头游戏是这样的，在游戏的开始，有几堆石头，两个选手轮流进行移动，一个合法的移动包括从其中一堆取走一块或多块石头，而不去移动其余的所有石头。不能进行合法移动的选手告负。(也可以规定取走最后一块石头的选手告负，因为不允许没有石头堆的局面。)图 9-22 所示的博弈树表示这种形式的给定开局的轮流取石头游戏，其中有 3 堆石头，分别包含 2 块、2 块和 1 块石头。用不同堆中石头数的无序表来表示每个局面(堆的顺序无关紧要)。第一个选手的初始移动可以导致 3 种可能的局面，因为这个选手可以从有 2 块石头的堆中取走 1 块石头(留下包含 1 块、1 块和 2 块石头的 3 堆)，可以从包含 2 块石头的堆中取走 2 块石头(留下包含 2 块和 1 块石头的 2 堆)，或者从包含 1 块石头的堆中取走 1 块石头(留下包含 2 块石头的 2 堆)。当只剩下包含 1 块石头的 1 堆时，就不可能进行合法移动了，所以这样的局面就是终局。由于轮流取石头游戏是非胜即负的游戏，所以用 +1 标记表示第一个选手获胜的终结顶点，用 −1 标记表示第二个选手获胜的终结顶点。 ■

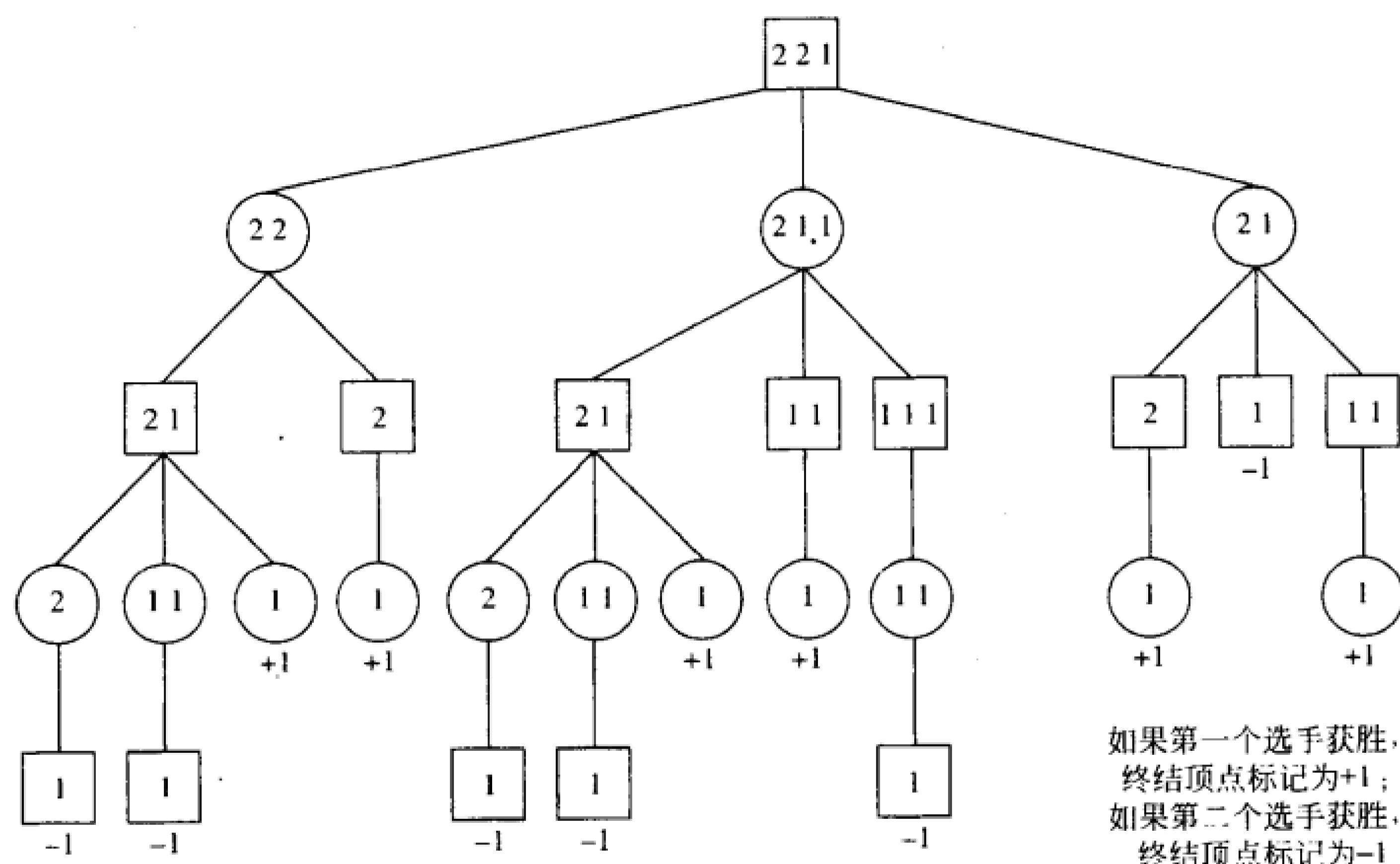


图 9-22 轮流取石头游戏的博弈树

例 6 圈叉游戏 圈叉游戏的博弈树非常大，在这里不能画出，尽管计算机能轻而易举地构造出这样的树。图 9-23a 显示圈叉游戏的博弈树的一部分。因为对称的顶点是等价的，所以只需要考虑图 9-23a 所示的 3 种可能的初始移动。还显示这个博弈树的导致图 9-23b 中终局的一个子树，其中一个能够获胜的选手进行了致胜的移动。 ■

可以用某种方式递归地定义博弈树中所有顶点的值，使得可以确定当两个选手都遵循最优策略时这个游戏的结果。所谓策略，就意味着一组规则，这些规则说明一个选手如何移动来赢得游戏。第一个选手的最优策略就是把这个选手的得分最大化的策略，第二个选手的最优策略就是把这个得分最小化的策略。现在递归地定义顶点的值。

定义 1 博弈树中顶点的值递归地定义为：

(i)一个树叶的值是当游戏在这个树叶所表示的局面里终止时第一个选手的得分。

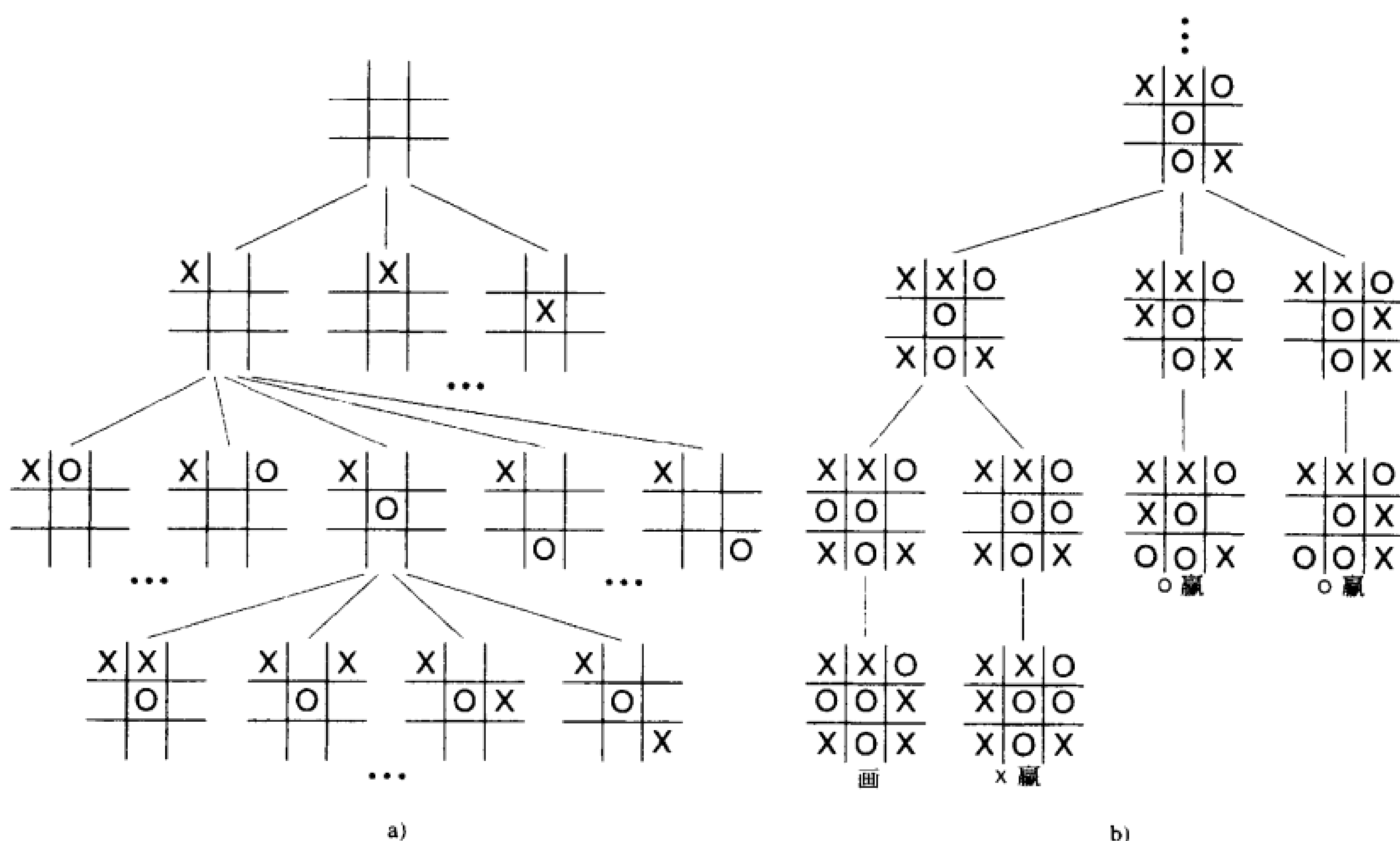


图 9-23 圈叉游戏的部分博弈树

(ii)偶数层内点的值是这个内点的子女的最大值，奇数层内点的值是这个内点的子女的最小值。

第一个选手移动到具有最大值的子女所表示的局面并且第二个选手移动到具有最小值的子女所表示的局面，这样的策略称为**最小最大策略**。当两个选手都遵循最小最大策略时，通过计算树根的值就可以确定谁将赢得游戏，这个值称为树的值。这是定理 3 的后果。

定理 3 如果两个选手都遵循最小最大策略并且从博弈树的某一个顶点所表示的局面开始进行游戏，则这个顶点的值说明第一个选手的得分。

证 将用归纳法来证明这个定理。

基础步骤：如果这个顶点是树叶，则根据定义指定给这个顶点的值就是第一个选手的得分。

归纳步骤：归纳假设一个顶点的子女的值就是第一个选手的得分，假定从这些顶点所表示的每一个局面中开始进行游戏。需要考虑两种情形，即当轮到第一个选手时和当轮到第二个选手时。

当轮到第一个选手时，这个选手遵循最小最大策略并且移动到具有最大值的子女所表示的局面。根据归纳假设，当从这个子女所表示的局面开始游戏并且遵循最小最大策略时，这个值就是第一个选手的得分。根据偶数层内点的值的定义的递归步骤(作为其子女的最大值)，当从这个顶点所表示的局面开始游戏时，这个顶点的值就是这个得分。

当轮到第二个选手时，这个选手遵循最小最大策略并且移动到具有最小值的子女所表示的局面。根据归纳假设，当从这个子女所表示的局面开始游戏并且遵循最小最大策略时，这个值就是第一个选手的得分。根据把奇数层内点的值作为其子女的最小值的递归定义，当从这个顶点所表示的局面开始游戏时，这个顶点的值就是这个得分。 □

注意 通过扩展定理 3 的证明，可以证明对于两个选手来说最小最大策略都是最优策略。

例 7 解释最小最大过程如何工作。它说明为例 5 的博弈树中的内点所指定的值。注意可以缩短所需要的计算，注意到对于非胜即负游戏来说，一旦找到方框顶点具有＋1 值的一个子女，则方框顶点的值也是＋1，因为＋1 是最大可能的得分。同样，一旦找到圆圈顶点具有－1 值的一个子女，则这个值也是这个圆圈顶点的值。

例 7 例 5 构造了具有包含 2 块、2 块和 1 块的 3 堆石头的开局的轮流取石头游戏的博弈树。图 9-24 说明这个博弈树的顶点的值。这些顶点的值是这样计算的：使用树叶的值并且每次向上计算 1 层。这个图的右边空白处说明究竟使用子女的最大值还是最小值来求出每层内点的值。例如，一旦求出了树根的 3 个子女的值，则这样求出树根的值：计算 max(1，－1，－1)＝1。由于根的值是 1，所以得出当两个选手都遵循最小最大策略时第一个选手获胜。■

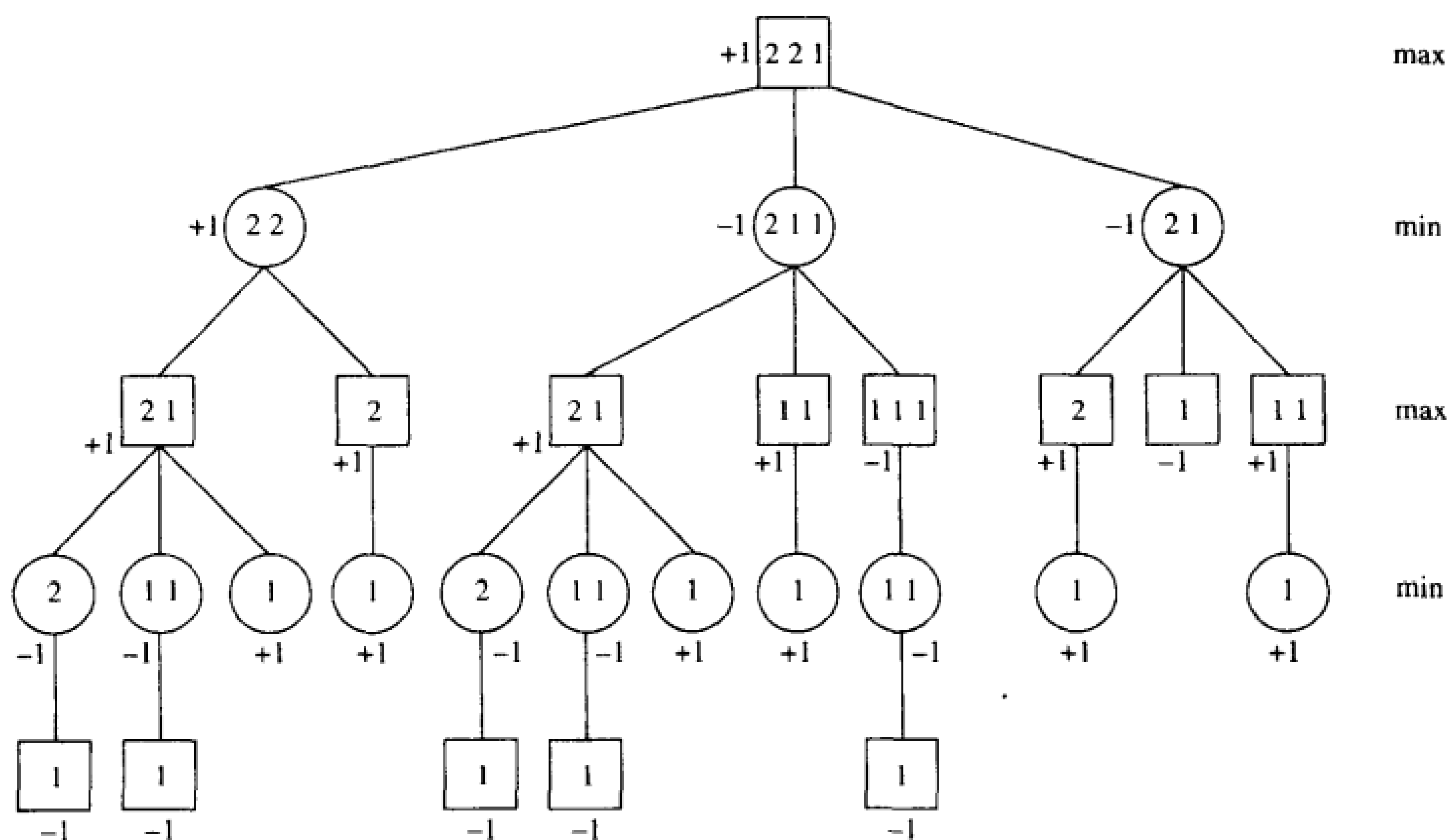


图 9-24 说明轮流取石头游戏中顶点的值

一些著名游戏的博弈树可能非同寻常地大，因为这些游戏有多种移动选择。例如，据估计象棋的博弈树有多达 10^{100} 个顶点！由于博弈树规模的原因，也许不可能直接使用定理 3 来研究这样的游戏，所以设计了各式各样的方法来帮助确定好的策略以及确定游戏的结果。一种被称作 $\alpha-\beta$ 剪枝的有用技巧减少了许多计算，它剪掉不能影响祖先顶点的值的那部分博弈树(关于 $\alpha-\beta$ 剪枝的信息，参考[Gr90])。另一种有用的方法是使用**求值函数**，当精确地计算博弈树中内点值不可行时，它就估计这些值。例如，在圈叉游戏中，可以使用不含圈 O(O 用来表示第二个选手的移动)的直行(行、列、对角线)数减去不含叉 X(X 用来表示第一个选手的移动)的直行数来作为一个局面的求值函数。这个求值函数给出了关于哪个选手在游戏中领先的一些说明。一旦插入求值函数的值，遵循最小最大策略使用的规则就可以计算出游戏的值。科学家研制出来的用来下棋的计算机程序，比如著名的深蓝程序，都是基于复杂的求值函数的。关于计算机如何下棋的更多信息请参看[Le91]。

练习

1. 用字母顺序建立下面这些单词的二叉搜索树：*banana*，*peach*，*apple*，*pear*，*coconut*，*mango* 和 *papaya*。

2. 用字母顺序建立下面这些单词的二叉搜索树：*oenology*，*phrenology*，*campanology*，*ornithology*，*ichthyology*，*limnology*，*alchemy* 和 *astrology*。
3. 为了在练习 1 的搜索树里找出下面每个单词的位置或者添加它们，而且每次都重新开始，分别需要多少次比较？

   a) *pear*　　b) *banana*　　c) *kumquat*　　d) *orange*
4. 为了在练习 2 的搜索树里找出下面每个单词的位置或者添加它们，而且每次都重新开始，分别需要多少次比较？

   a) *palmistry*　　b) *etymology*　　c) *paleontology*　　d) *glaciology*
5. 用字母顺序构造下面句子里的单词的二叉搜索树："*The quick brown fox jumps over the lazy dog*"。
6. 为了在四枚硬币中找出一枚较轻的伪币，需要用天平称多少次？描述用这些次数的称重来找出较轻的伪币的算法。
7. 若一枚伪币与其他硬币质量不等，那么为了在 4 枚硬币中找出这枚伪币，需要用天平称多少次？描述用同样的称重次数来找出这枚伪币的算法。

*8. 若一枚伪币与其他硬币质量不等，那么为了在 8 枚硬币中找出这枚伪币，需要用天平称多少次？描述用同样的称重次数来找出这枚伪币的算法。

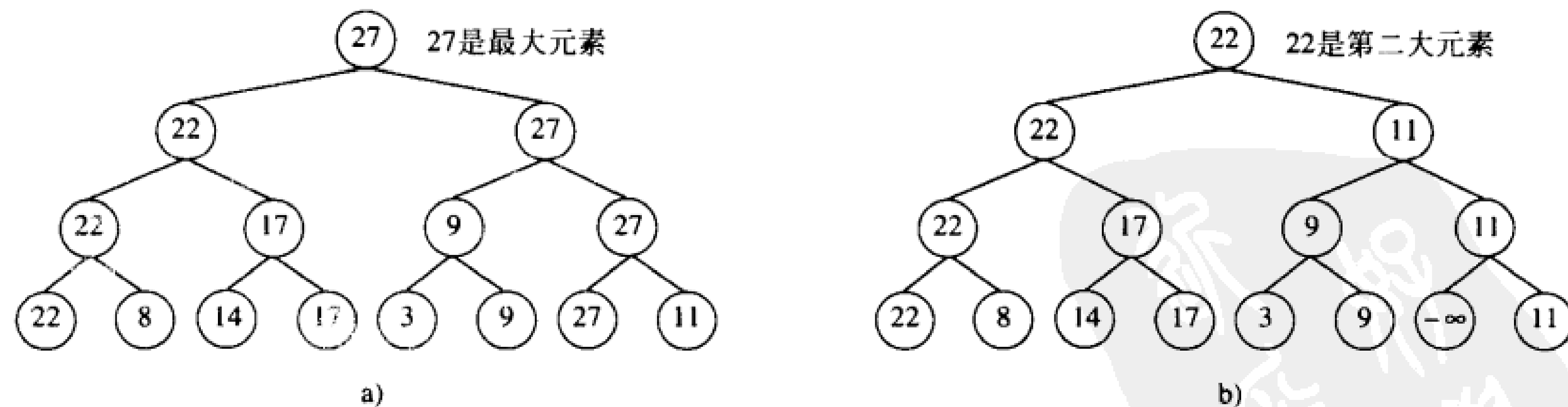
*9. 若一枚伪币比其他硬币较轻，那么为了在 12 枚硬币中找出这枚伪币，需要用天平称多少次？描述用同样的称重次数来找出这枚伪币的算法。

*10. 四枚硬币中一枚可能是伪币。伪币与其他硬币质量不等。那么为了确定是否有一个伪币，以及若有伪币，确定它是比其他硬币较重还是较轻，使用一台天平称，需要称多少次？描述用同样的称重次数来找出这枚伪币并且确定它是较轻还是较重的算法。

11. 求排序 4 个元素所需要的最少比较次数并且设计一个能够依此次数实现的算法。

*12. 求排序 5 个元素所需要的最少比较次数并且设计一个能够依此次数实现的算法。

竞赛图排序是通过构造有序二叉树来工作的排序算法。用将成为树叶的顶点来表示待排序的元素。就如同构造表示循环赛比赛胜者的树那样，一次构造这个树的一层。从左向右，比较成对的相邻元素，加入用所比较的两个元素中较大的那个来标记的一个父母顶点。在每一层在顶点的标记之间进行类似的比较，直到到达了用最大元素标记的树根为止。22，8，14，17，3，9，27，11 的竞赛图排序所构造的树如下图(a)所示。一旦确定了最大元素，具有这个标记的树叶就重新标记为$-\infty$，定义为比每个元素都小。从这个顶点直到树根的路径上所有顶点的标记都重新计算，如下图(b)所示。这样就产生了第二大元素。这个过程继续进行下去，直到整个表都已经排序为止。



13. 完成表 22，8，14，17，3，9，27，11 的竞赛图排序。说明在每个步骤上顶点的标记。
14. 用竞赛图排序来排序表 17，4，1，5，13，10，14，6。
15. 用伪码描述竞赛图排序。
16. 假设对于某个正整数 k 来说，待排序元素的个数 n 等于 2^k，使用竞赛图排序来求这个表的最大元素，确定竞赛图排序所用的比较次数。
17. 竞赛图排序求第二大元素、第三大元素……，直到第$(n-1)$大(或第二小)元素所使用的比较次数是

多少？

18. 证明：竞赛图排序需要 $\Theta(n\log n)$ 次比较来排序 n 个元素的表。（提示：假设对于某个正整数 k 来说，$n=2^k$，插入适当数目的哑元，比如 $-\infty$，定义它们比所有整数都小。）
19. 下面哪些编码是前缀码？

    a) a：11，e：00，t：10，s：01

    b) a：0，e：1，t：01，s：001

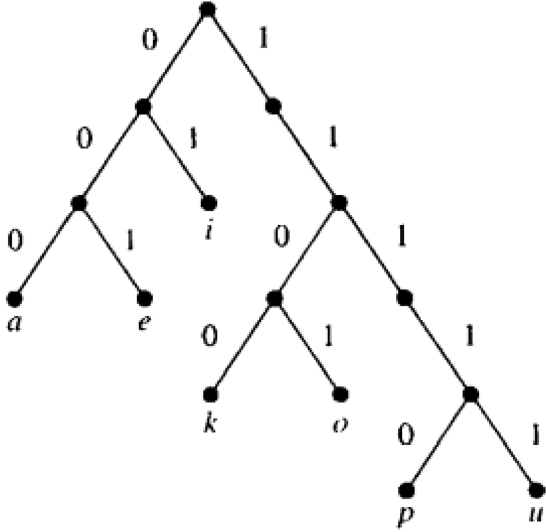
    c) a：101，e：11，t：001，s：011，n：010

    d) a：010，e：11，t：011，s：1011，n：1001，i：10101
20. 构造表示下面编码方案的前缀码的二叉树。

    a) a：11，e：0，t：101，s：100

    b) a：1，e：01，t：001，s：0001，n：00001

    c) a：1010，e：0，t：11，s：1011，n：1001，i：10001
21. 若编码方案是用右面的树来表示，那么什么是 a、e、i、k、o、p 和 u 的编码？



22. 给定编码方案 a：001，b：0001，e：1，r：0000，s：0100，t：011，x：01010，找出用下面的位串来表示的单词。

    a) 01110100011　　b) 0001110000

    c) 0100101010　　d) 01100101010
23. 用赫夫曼编码来编码具有给定频率的如下符号：a：0.20，b：0.10，c：0.15，d：0.25，e：0.30。编码一个符号所需要的平均位数是多少？
24. 用赫夫曼编码来编码具有给定频率的如下符号：A：0.10，B：0.25，C：0.05，D：0.15，E：0.30，F：0.07，G：0.08。编码一个符号所需要的平均位数是多少？
25. 为如下符号和频率构造两个不同的赫夫曼编码：t：0.2，u：0.3，v：0.2，w：0.3。
26. a) 以两种不同的方式用赫夫曼编码来编码具有这些频率的符号：a：0.4，b：0.2，c：0.2，d：0.1，e：0.1，在算法中用不同的方式打破平局。首先，在算法的每个阶段从权最小的树中选择顶点数最多的两个树来组合。其次，在每个阶段从权最小的树中选择顶点数最少的两个树来组合。

    b) 计算用每种编码来编码一个符号所需要的平均位数并且对每种编码计算这个位数的方差。对于编码一个符号所需要的位数的方差，哪种打破平局的过程所产生的会小一些？
27. 为英文字母表的字母构造赫夫曼编码，其中典型英文文本中字母的频率如下表所示。

字母	频率	字母	频率
A	0.0817	N	0.0662
B	0.0145	O	0.0781
C	0.0248	P	0.0156
D	0.0431	Q	0.0009
E	0.1232	R	0.0572
F	0.0209	S	0.0628
G	0.0182	T	0.0905
H	0.0668	U	0.0304
I	0.0689	V	0.0102
J	0.0010	W	0.0264
K	0.0080	X	0.0015
L	0.0397	Y	0.0211
M	0.0277	Z	0.0005

假设 m 是正整数且 $m \geqslant 2$。对于 N 个符号的集合来说，类似于二元赫夫曼编码的构造，可以构造 m 元赫夫曼编码。在初始步骤，把由 $((N-1)\mathbf{mod}(m-1))+1$ 个权最小的单个顶点所组成的树组合成以这些顶点作为树叶的一棵树。在每个后续步骤，把权最小的 m 棵树组合成一棵 m 元树。

28. 用伪码描述 m 元赫夫曼编码算法。

29. 使用符号 0、1 和 2，用三元($m=3$)赫夫曼编码来编码具有给定频率的这些字母：A：0.25，E：0.30，N：0.10，R：0.05，T：0.12，Z：0.18。

30. 考虑具有频率 A：0.80，B：0.19，C：0.01 的三个符号 A、B 和 C。
  a)为这三个符号构造赫夫曼编码。
  b)通过把两个符号的块 AA、AB、AC、BA、BB、BC、CA、CB 和 CC 分组，形成 9 个符号的一个新的集合。为这 9 个符号构造赫夫曼编码，假设在原始文本中符号的出现是独立的。
  c)比较一下(a)中三个符号的赫夫曼编码与(b)构造的 9 个双符号块的赫夫曼编码来编码文本所需要的平均位数。哪一种编码方式更有效？

31. 给定 $n+1$ 个符号 x_1，x_2，…，x_n，x_{n+1}，它们在一个符号串中分别出现 1，f_1，f_2，…，f_n 次，其中 f_j 是第 j 个斐波那契数。当在赫夫曼编码算法的每个阶段考虑所有可能的打破平局的选择时，用来编码一个符号的最大位数是多少？

*32. 证明：对于所有的二叉前缀码来说，如果使用最少的位来表示一个符号串，赫夫曼编码是最优的。

33. 画出轮流取石头游戏的博弈树，假设开局包括分别有 2 块和 3 块石头的两堆石头。在画这棵树的时候，用同一个顶点表示相同移动所导致的对称局面。求出这个博弈树每个顶点的值。如果两个选手都遵循最优策略，则哪个选手获胜？

34. 画出轮流取石头游戏的博弈树，假设开局包括分别有 1 块、2 块和 3 块石头的三堆石头。在画这棵树的时候，用同一个顶点表示相同移动所导致的对称局面。求出这棵博弈树每个顶点的值。如果两个选手都遵循最优策略，则哪个选手获胜？

35. 假设在轮流取石头游戏中修改获胜选手的得分，使得当 n 是到达终局前所做合法移动的步数时得分就是 n 美元。求第一个选手的得分，假设开局包括：
  a)分别有 1 块和 3 块石头的两堆石头。
  b)分别有 2 块和 4 块石头的两堆石头。
  c)分别有 1 块、2 块和 3 块石头的三堆石头。

36. 假设在轮流取石头游戏的一个变种中，允许一个选手要么从一堆取走 1 块以上的石头，要么把两堆石头合并成一堆，只要至少还剩下一堆石头。画出这个游戏变种的博弈树，假设开局由分别含有 2 块、2 块和 1 块石头的 3 堆石头组成。求出这棵博弈树中每个顶点的值，并且确定当两个选手都遵循最优策略时哪个选手获胜。

37. 画出圈叉游戏博弈树从下列每个局面开始的子树。确定每个子树的值。

a)

O	X	X
X	O	O
		X

b)

X	O	X
O	X	X
		O

c)

X		O
O	O	
X		X

d)

	O	X
	X	O
X	O	

38. 假设圈叉游戏的前 4 步移动如下图所示。第一个选手(用×标记其移动)是否总能获胜？

a)

X	O	X
	O	

b)

X	O	X
	O	

c)

	O	X
	X	
O		

d)

X		
X	O	
O		

39. 证明：如果轮流取石头游戏从包含相同数目的2堆石头开始，只要这个数目至少是2，则当两个选手都遵循最优策略时第二个选手获胜。
40. 证明：如果轮流取石头游戏从包含不同数目的2堆石头开始，则当两个选手都遵循最优策略时第一个选手获胜。
41. 跳棋博弈树的根有多少个子女？有多少个孙子女？
42. 轮流取石头游戏博弈树的根有多少个子女？有多少个孙子女？假设开局是：
    a)分别有4块和5块石头的堆。
    b)分别有2块、3块和4块石头的堆。
    c)分别有1块、2块、3块和4块石头的堆。
    d)分别有2块、2块、3块、3块和5块石头的堆。
43. 画出圈叉游戏博弈树前两步移动所对应的层。指明正文中所提到的求值函数的值，这个函数给局面指定不含○的直行数减去不含×的直行数来作为这一层每个顶点的值，并且在求值函数给出这些顶点的正确值的假设下，对这些顶点计算树的值。
44. 用伪码描述当两个选手都遵循最小最大策略时确定博弈树的值的算法。

9.3 树的遍历

9.3.1 引言

有序根树常常用来保存信息。掌握一些访问有序根树的每个顶点以存取数据的算法是非常必要的。几个重要的算法将在下面描述。有序根树也可以用来表示各种类型的表达式，比如由数字、变量和运算所组成的算术表达式。对用来表示这些表达式的有序根树来说，它的顶点的一些不同的列表在这些表达式的求值中很有用。

9.3.2 通用地址系统

遍历有序根树的所有顶点的过程，都依赖于子女的顺序。在有序根树里，一个内点的子女从左向右地显示在表示这些有向图的图形里。

下面将描述一种完全地排序有序根树顶点的方法。为了产生这个顺序，必须首先标记所有的顶点。如下递归地完成这件事。

1. 用整数0标记根。然后用1，2，3，…，k从左向右标记它的k个子女(在1层上)。

2. 对在n层上带标记A的每个顶点v，按照从左向右画出它的k_v个子女的顺序，用$A.1$，$A.2$，…，$A.k$标记它的k_v个子女。

遵循这个过程，对$n\geqslant 1$来说，在n层上的顶点v标记成$x_1.x_2.\cdots.x_n$，其中从根到v的唯一通路经过1层的第x_1个顶点，以及2层的第x_2个顶点，依此类推。这样的标记称为根树的**通用地址系统**。

可以利用顶点在通用地址系统里的标记的字典顺序将这些顶点完全排序。若存在$i(0\leqslant i\leqslant n)$满足$x_1=y_1$，$x_2=y_2$，…，$x_{i-1}=y_{i-1}$，并且$x_i<y_i$；或者若$n<m$并且对$i=1$，2，…，$n$来说$x_i=y_i$，那么标记着$x_1.x_2.\cdots.x_n$的顶点就小于标记着$y_1.y_2.\cdots.y_n$的顶点。

例1 在如图9-25所示的有序根树的顶点的旁边，显示出了通用地址系统的标记。这些标记的字典顺序是

$$0<1<1.1<1.2<1.3<2<3<3.1<3.1.1<3.1.2<3.1.2.1<3.1.2.2<3.1.2.3<3.1.2.4<3.1.3<3.2<4<4.1<5<5.1<5.1.1<5.2<5.3$$ ■

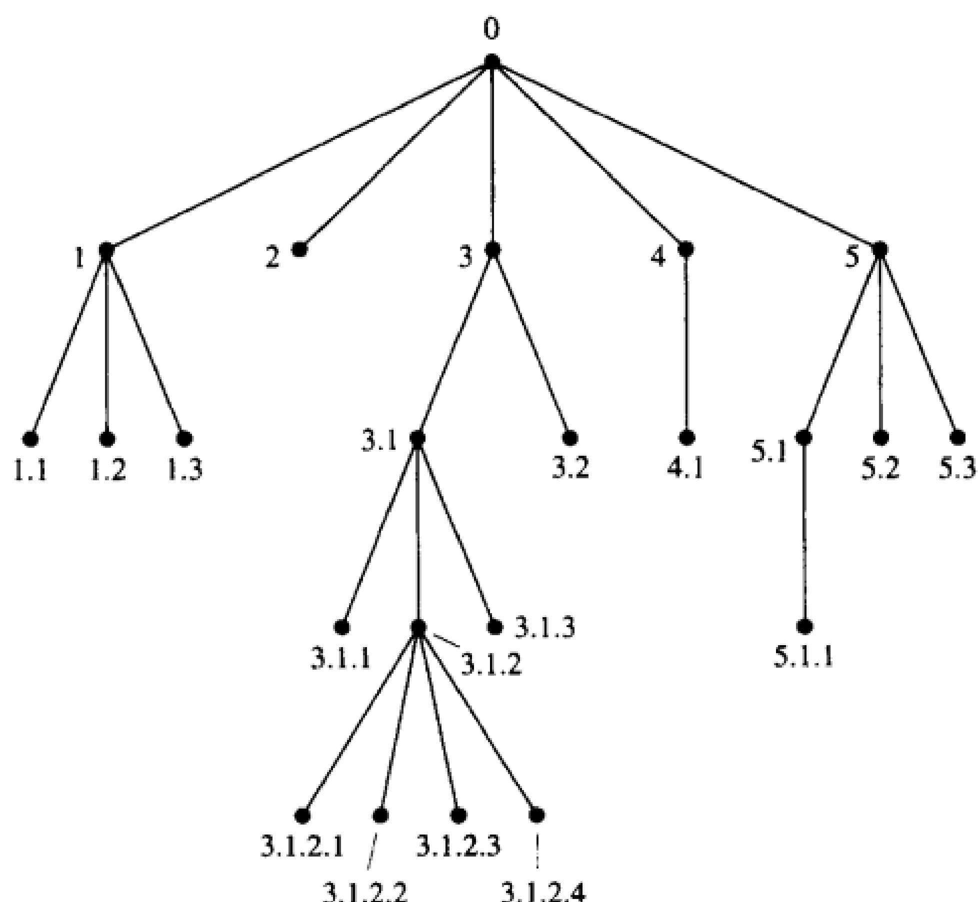


图 9-25 有序根树的通用地址系统

9.3.3 遍历算法

系统地访问有序根树每个顶点的过程都称为*遍历算法*。下面描述三个最常用的此类算法：*前序遍历*、*中序遍历*和*后序遍历*。这些算法都可以递归地定义。首先定义前序遍历。

定义 1 设 T 是带根 r 的有序根树。若 T 只包含 r，则 r 是 T 的*前序遍历*。否则，假定 T_1，T_2，…，T_n 是 T 里在 r 处从左向右的子树。*前序遍历*首先访问 r。它接着以前序来遍历 T_1，然后以前序来遍历 T_2，依此类推，直到以前序来遍历了 T_n 为止。

读者应当验证，有序根树的前序遍历给出与利用通用地址系统所得出的顺序相同的顶点顺序。图 9-26 说明如何执行前序遍历。

例 2 说明前序遍历。

例 2 前序遍历以什么顺序访问图 9-27 所示的有序根树里的顶点？

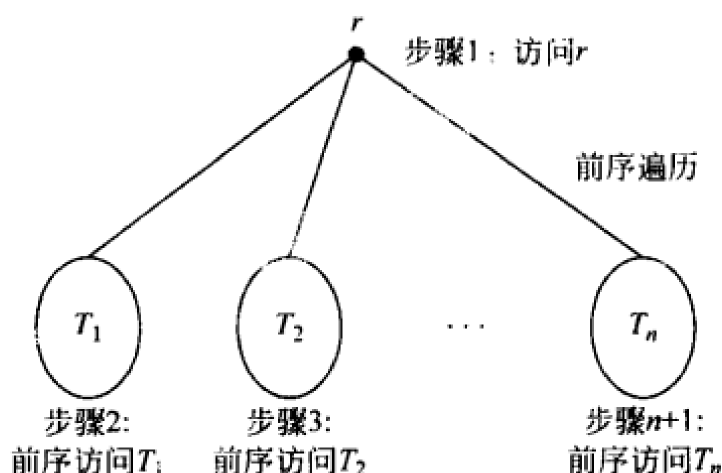


图 9-26 前序遍历

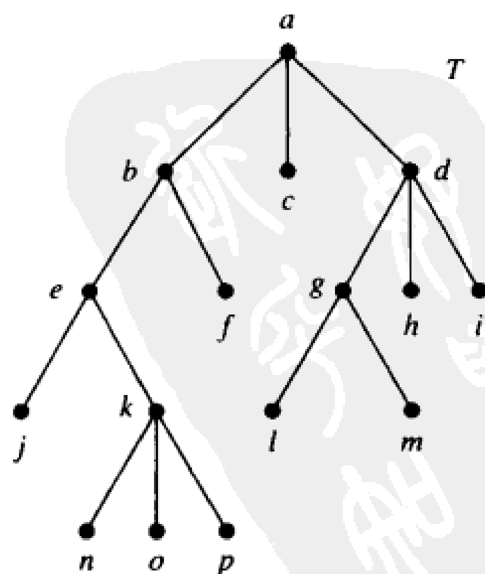


图 9-27 有序根树 T

解 T 的前序遍历的步骤如图 9-28 所示。这样以前序来遍历 T，首先列出根 a，接着依次是带根 b 的子树的前序列表，带根 c 的子树(它只有 c)的前序列表和带根 d 的子树的前序列表。

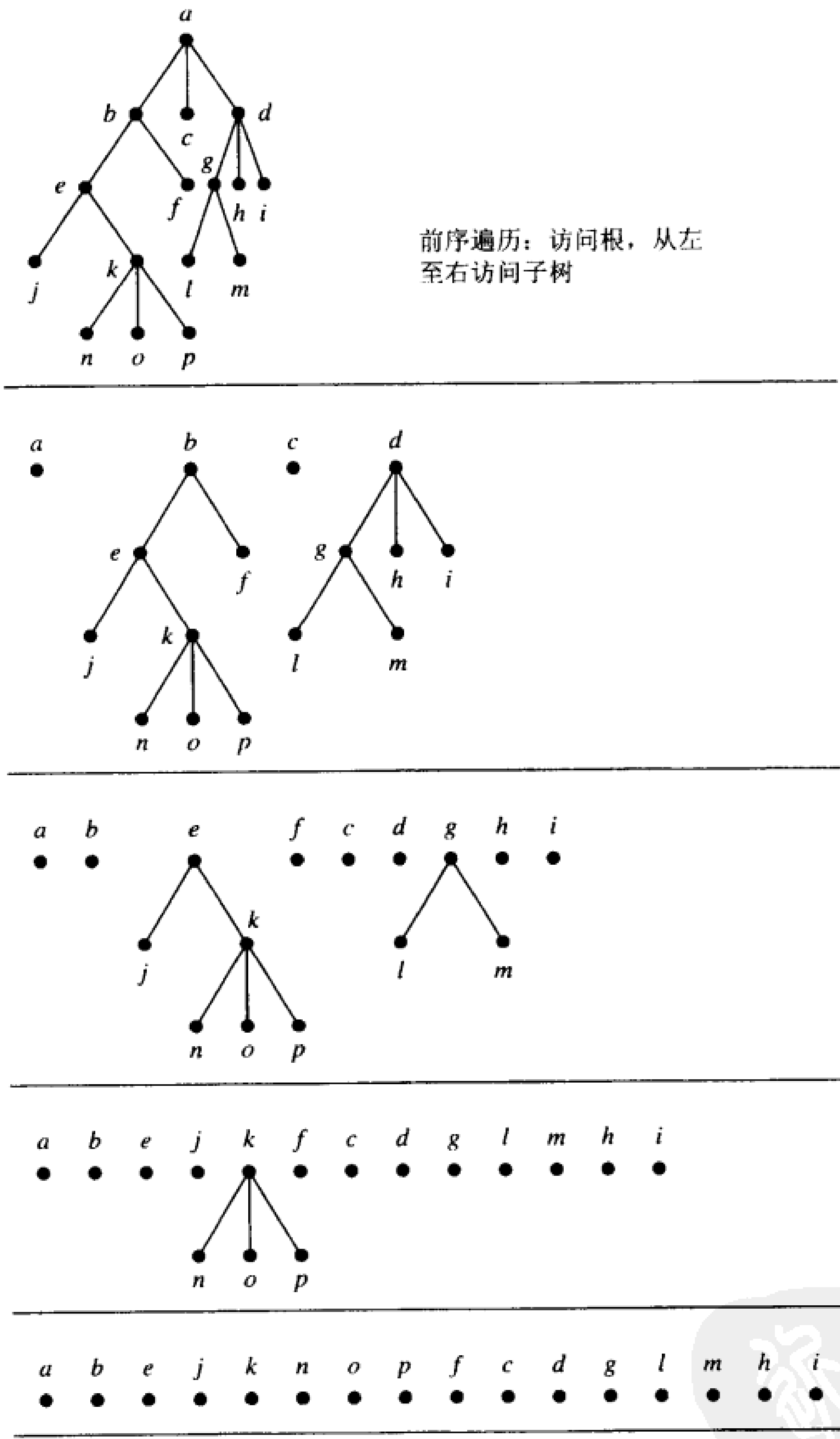


图 9-28 T 的前序遍历

带根 b 的子树的前序列表首先列出 b，然后以前序列出带根 e 的子树的顶点，然后以前序列出带根 f 的子树(它只有 f)的顶点。带根 d 的子树的前序列表首先列出 d，接着是带根 g 的子树的前序列表，接着是带根 h 的子树(它只有 h)，接着是带根 i 的子树(它只有 i)。

带根 e 的子树的前序列表首先列出 e，接着是带根 j 的子树(它只有 j)的前序列表，接着是带根 k 的子树的前序列表。带根 g 的子树的前序列表是 g 接着 l，接着是 m。带根 k 的子树的前序列表

是 k, n, o, p。所以，T 的前序遍历是 a, b, e, j, k, n, o, p, f, c, d, g, l, m, h, i。 ■

现在将定义中序遍历。

定义 2 设 T 是带根 r 的有序根树。若 T 只包含 r，则 r 是 T 的*中序遍历*。否则，假定 T_1, T_2, …, T_n 是 T 里在 r 处从左向右的子树。*中序遍历*首先以前序来遍历 T_1，然后访问 r。它接着以中序来遍历 T_2，中序遍历 T_3，依此类推，直到以中序来遍历了 T_n 为止。

图 9-29 说明如何执行中序遍历。例 3 说明对一棵特定的树，如何执行中序遍历。

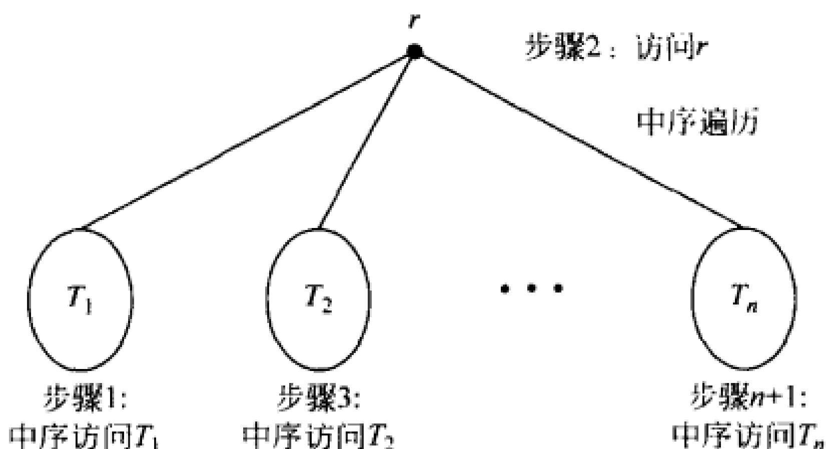


图 9-29 中序遍历

例 3 中序遍历以什么顺序访问图 9-27 所示的有序根树 T 里的顶点？

Extra Examples **解** T 的中序遍历的步骤显示在图 9-30 里。中序遍历首先是带根 b 的子树的中序遍历，然后是根 a，带根 c 的子树（它只有 c）的中序列表和带根 d 的子树的中序列表。

带根 b 的子树的中序列表，首先是带根 e 的子树的中序列表，然后是根 b，以及 f。带根 d 的子树的中序列表，首先是带根 g 的子树的中序列表，接着是根 d，接着是 h，接着是 i。

带根 e 的子树的中序列表是 j，接着是根 e，接着是带根 k 的子树的中序列表。带根 g 的子树的中序列表是 l, g, m。带根 k 的子树的中序列表是 n, k, o, p。所以，这个根树的中序遍历是 j, e, n, k, o, p, b, f, a, c, l, g, m, d, h, i。 ■

现在定义后序遍历。

定义 3 设 T 是带根 r 的有序根树。若 T 只包含 r，则 r 是 T 的*后序遍历*。否则，假定 T_1, T_2, …, T_n 是 T 里在 r 处从左向右的子树。*后序遍历*首先以后序来遍历 T_1，然后以后序来遍历 T_2，……然后以后序来遍历 T_n，最后访问 r。

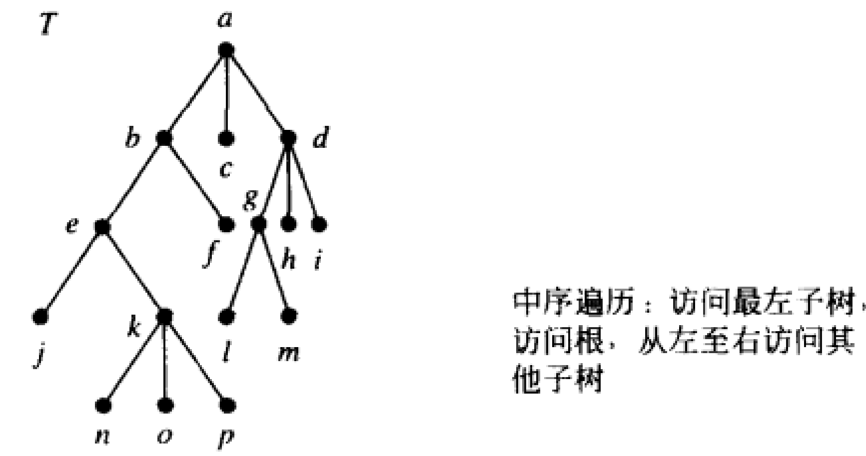
图 9-31 说明后序遍历是如何执行的。例 4 说明后序遍历如何工作。

例 4 后序遍历以什么顺序访问图 9-27 所示有序根树 T 里的顶点？

Extra Examples **解** T 的后序遍历的步骤显示在图 9-32 里。后序遍历首先是带根 b 的子树的后序遍历，然后是带根 c 的子树（它只有 c）的后序遍历，带根 d 的子树的后序遍历，接着是根 a。

带根 b 的子树的后序遍历首先是带根 e 的子树的后序遍历，接着是 f，接着是根 b。带根 d 的子树的后序遍历首先是带根 g 的子树的后序遍历，接着是 h，接着是 i，接着是根 d。

带根 e 的子树的后序遍历是 j，接着是带根 k 的子树的后序遍历，接着是根 e。带根 g 的子树的后序遍历是 l, m, g。带根 k 的子树的后序遍历是 n, o, p, k。因此，根树 T 的后序遍历是 j, n, o, p, k, e, f, b, c, l, m, g, h, i, d, a。 ■



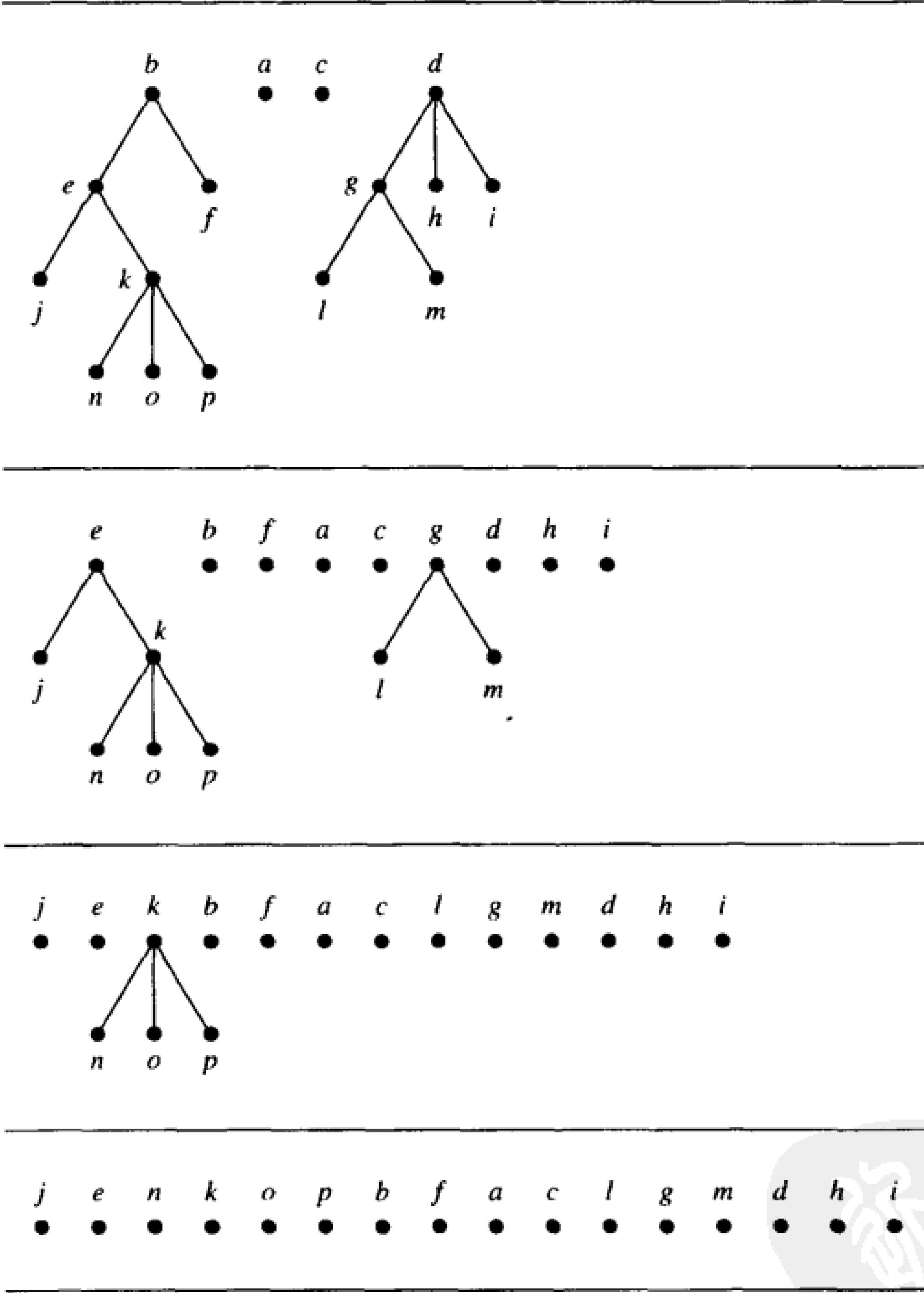


图 9-30 T 的中序遍历

有一些简易的方法以前序、中序和后序来列出有序根树的顶点。首先从根开始围绕有序根树画一条曲线，如图 9-33 里的例子所示，沿着边移动。可以这样按照前序来列出顶点：当曲线第一次经过一个顶点时，就列出这个顶点。可以这样按照中序来列出顶点：当曲线第一次经过一个树叶时，就列出这个树叶，当曲线第二次经过一个内点时就列出这个内点。可以这样按照后序来列出顶点：当曲线最后一次经过一个顶点而返回这个顶点的父母时，就列出这个顶点。当在图 9-33

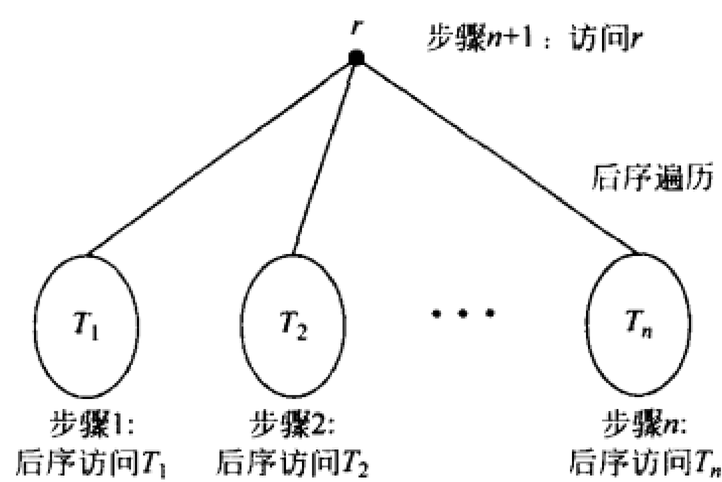


图 9-31　后序遍历

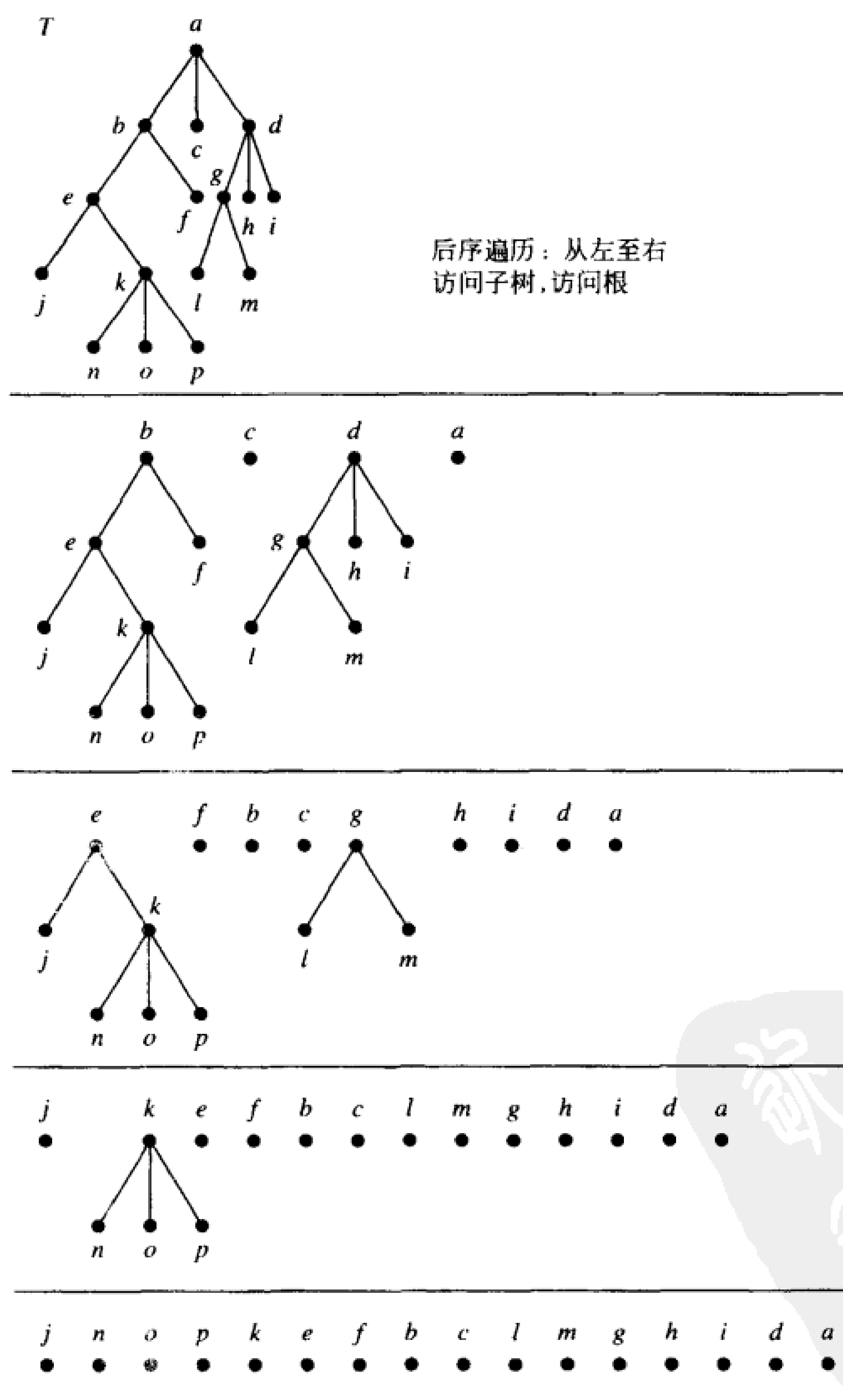


图 9-32　T 的后序遍历

的根树里这样做时，结果是前序遍历给出 a，b，d，h，e，i，j，c，f，g，k；中序遍历给出 h，d，b，i，e，j，a，f，c，k，g；后序遍历给出 h、d，i，j，e，b，f，k，g，c，a。

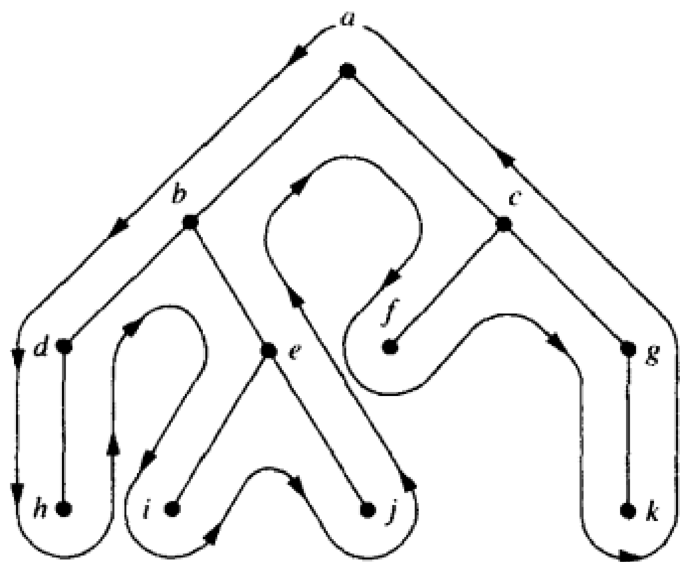


图 9-33 以前序、中序和后序来遍历有序根树的快捷方法

这些以前序、中序和后序来遍历有序根树的算法，最容易用递归来表示。

算法 1 前序遍历

procedure *preorder*(*T*：有序根树)
r := *T* 的根
列出 *r*
for 从左到右的 *r* 的每个子女 *c*
begin
  T(*c*) := 以 *c* 为根的子树
  preorder(*T*(*c*))
end

算法 2 中序遍历

procedure *inorder*(*T*：有序根树)
r := *T* 的根
if *r* 是树叶 **then** 列出 *r*
else
begin
  l := 从左到右的 *r* 的第一个子女
  T(*l*) := 以 *l* 为根的子树
  inorder(*T*(*l*))
  列出 *r*
  for 除 *l* 外从左到右的 *r* 的每个子女 *c*
    T(*c*) := 以 *c* 为根的子树
    inorder(*T*(*c*))
end

```
算法 3  后序遍历
procedure postorder(T: 有序根树)
r := T 的根
for 从左到右的 r 的每个子女 c
begin
  T(c) := 以 c 为根的子树
  postorder(T(c))
end
列出 r
```

9.3.4 中缀、前缀和后缀记法

可以用有序树来表示复杂的表达式，比如复合命题，集合的组合，以及算术表达式。例如，考虑由运算＋(加)、－(减)、*(乘)、/(除)、↑(幂)所组成的算术表达式的表示。我们将用括号来说明运算次序。有序根树可以用来表示这样的表达式，其中内点表示运算，树叶表示变量或数字。每个运算都作用在它的左子树和右子树上(以此顺序)。

例 5 表示表达式$((x+y)\uparrow 2)+((x-4)/3)$的有序根树是什么？

解 这个表达式的二叉树可以自底向上来构造。首先，构造表达式 $x+y$ 的子树。然后加入这个子树作为表示$(x+y)\uparrow 2$ 的更大子树的一部分。同样，构造表达式 $x-4$ 的子树，然后加入这个子树到表示$(x-4)/3$ 的子树里。最后组合表示$(x+y)\uparrow 2$ 与$(x-4)/3$ 的子树来形成表示$((x+y)\uparrow 2)+((x-4)/3)$的有序根树。这些步骤都显示在图 9-34 里。■

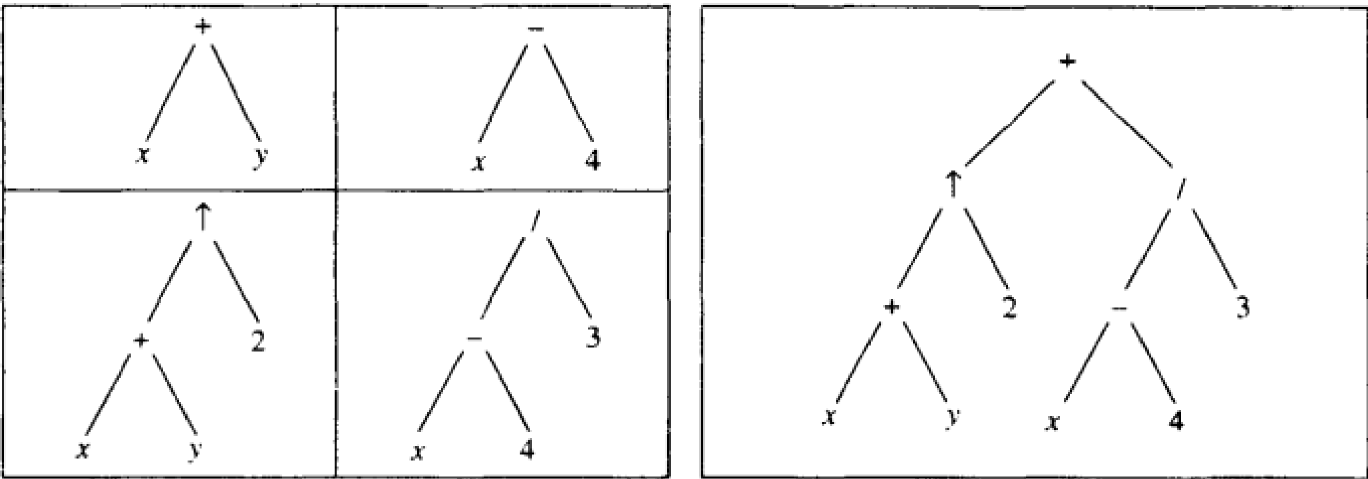


图 9-34　表示$((x+y)\uparrow 2)+((x-4)/3)$的二叉树

对表示一个表达式的二叉树的中序遍历，产生原来的表达式，其中元素和运算都是按原有的次序出现，例外的是一元运算，它们紧随运算对象。例如，图 9-35 里的二叉树分别表示表达

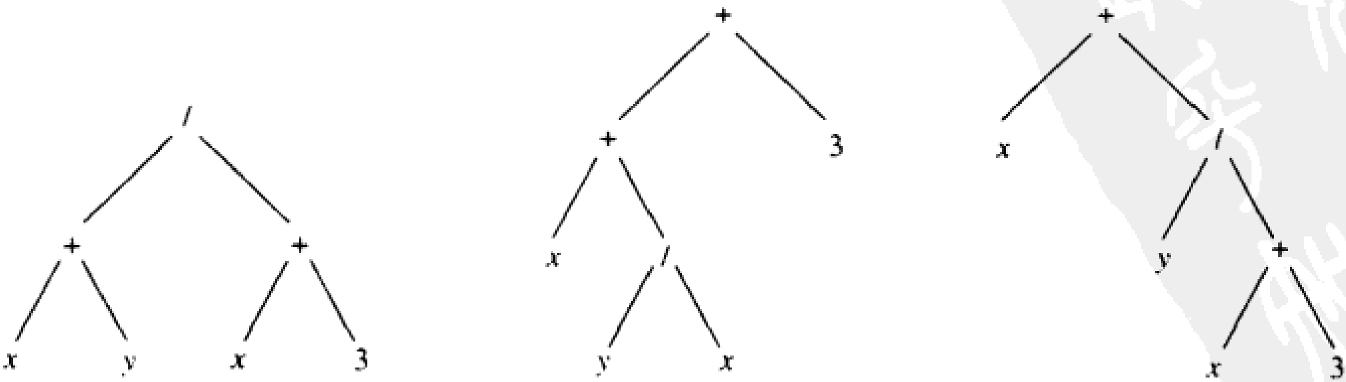


图 9-35　表示$(x+y)/(x+3)$，$(x+(y/x))+3$ 和 $x+(y/(x+3))$的根树

式$(x+y)/(x+3)$，$(x+(y/x))+3$ 和 $x+(y/(x+3))$，对它们的中序遍历都得出中缀表达式 $x+y/x+3$。为了让这样的表达式无二义性，每当遇到运算时，就有必要在中序遍历里包含括号。以这种方式获得的带完整括号的表达式称为*中缀形式*。

当以前序来遍历表达式的根树时，就获得它的*前缀形式*。写成前缀形式的表达式称为*波兰记法*，它的命名由来于逻辑学家扬·武卡谢维奇[⊖]（其实他是乌克兰人而非波兰人）。前缀记法下的表达式（其中每个运算都有规定的运算对象数）都是无二义性的。对这个事实的验证留给读者作为练习。

例 6 $((x+y)\uparrow 2)+((x-4)/3)$的前缀形式是什么？

解 通过遍历图 9-34 所示的表示这个表达式的二叉树，就获得它的前缀形式。这样就产生$+\uparrow+xy2/-x43$。■

在表达式的前缀形式里，二元运算（比如＋）在它的两个运算对象之前。因此，可以从右向左地求前缀形式的表达式的值。当遇到一个运算时，就对在这个运算右边紧接着的两个运算对象来执行相应的运算。另外，每当一个运算执行时，就认为结果是新的运算对象。

例 7 前缀表达式$+-*235/\uparrow 234$的值是什么？

解 用来从右向左地求这个表达式的值所用的步骤，以及用右边的运算对象来执行的运算，都如图 9-36 所示。这个表达式的值是 3。■

通过以后序来遍历表达式的二叉树，就获得它的*后缀形式*。写成后缀形式的表达式称为是*逆波兰记法*。逆波兰记法下的表达式都是无二义性的，所以就不需要括号。对这个事实的验证留给读者。

例 8 $((x+y)\uparrow 2)+((x-4)/3)$的后缀形式是什么？

解 这个表达式的后缀形式是这样获得的：执行图 9-34 所示的表示它的二叉树的后序遍历，这样就产生后缀表达式$x\ y+2\ \uparrow\ x\ 4-3\ /+$。■

在表达式的后缀形式里，二元运算都是在它的两个运算对象之后。所以，为了从一个表达式的后缀形式来求它的值，就从左向右地进行，每当一个运算后面跟着两个运算对象时，就执行这个运算。在一个运算执行之后，这个运算的结果就成为一个新的运算对象。

例 9 后缀表达式$7\ 2\ 3\ *-4\ \uparrow\ 9\ 3\ /+$的值是什么？

解 如图 9-37 所示，求这个表达式的值所用的步骤是这样的：从左边开始，当两个运算对象后面接着一个运算时，就执行这个运算。这个表达式的值是 4。■

⊖ **扬·武卡谢维奇(Jan Łukasiewicz，1878—1956)** 武卡谢维奇出生在勒沃的一个波兰语家庭。在他出生时勒沃是奥地利的一部分，现在则属于乌克兰。他的父亲是奥地利军队的上尉。武卡谢维奇在高中时开始对数学感兴趣。他在勒沃大学的本科和研究生阶段兼修数学和哲学。在完成博士工作后他成为那里的讲师，并且在1911 年被任命为教授。当华沙大学在 1915 年重新开办为波兰大学时，武卡谢维奇接受邀请成为教员。在 1919 年他担任波兰教育大臣。从 1920 年到 1939 年他回到华沙大学的教授职位上并一直呆在那里，曾两度担任这所大学的校长。

武卡谢维奇是著名的逻辑学华沙学派的共同发起人之一。他在 1928 年出版了著名的教材《数理逻辑基础》(*Elements of Mathematical Logic*)。受他的影响，数理逻辑在波兰成为数学和科学本科生的必修课程。他的讲座非常出色，甚至吸引了人文学科的学生。

武卡谢维奇和他的妻子在第二次世界大战期间遭受了极大的迫害，在他死后出版的自传中记录了这些迫害。战后他们在比利时过着流放生活。幸运的是，在 1949 年他获得在都柏林的爱尔兰皇家科学院的职务。

武卡谢维奇一生致力于数理逻辑的研究。他关于三值逻辑的工作是对这个题目的重要贡献。不过，在数学和计算机科学界他最著名的，是他引入了无括号记法，现在称为波兰记法。

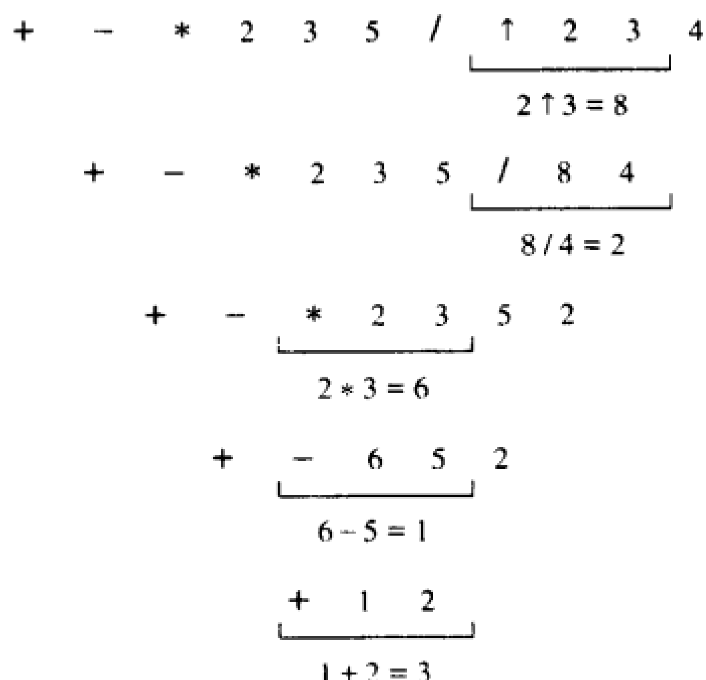


图 9-36 求一个前缀表达式的值

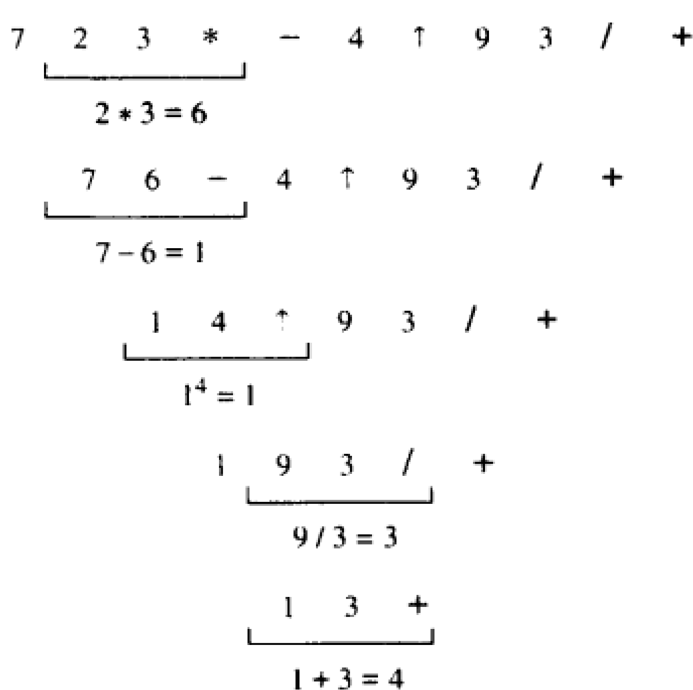


图 9-37 求一个后缀表达式的值

根树可以用来表示其他类型的表达式，比如那些表示复合命题、集合组合的表达式。在这些例子里出现比如命题否定这样的一元运算。为了表示这样的运算及其运算对象，就用顶点表示运算并且用这个顶点的子女表示运算对象。

例 10 求表示复合命题$(\neg(p\wedge q))\leftrightarrow(\neg p\vee\neg q)$的有序根树。然后用这个根树求这个表达式的前缀、后缀和中缀形式。

Extra Examples **解** 这个复合命题的有序根树是自底向上地构造的。首先，构造$\neg p$和$\neg q$的子树(其中把$\neg$当作一元运算)。另外，构造$p\wedge q$的子树。然后构造$\neg(p\wedge q)$和$(\neg p)\vee(\neg q)$的子树。最后，用这两个子树来构造最终的根树。这个过程的步骤显示在图 9-38 中。

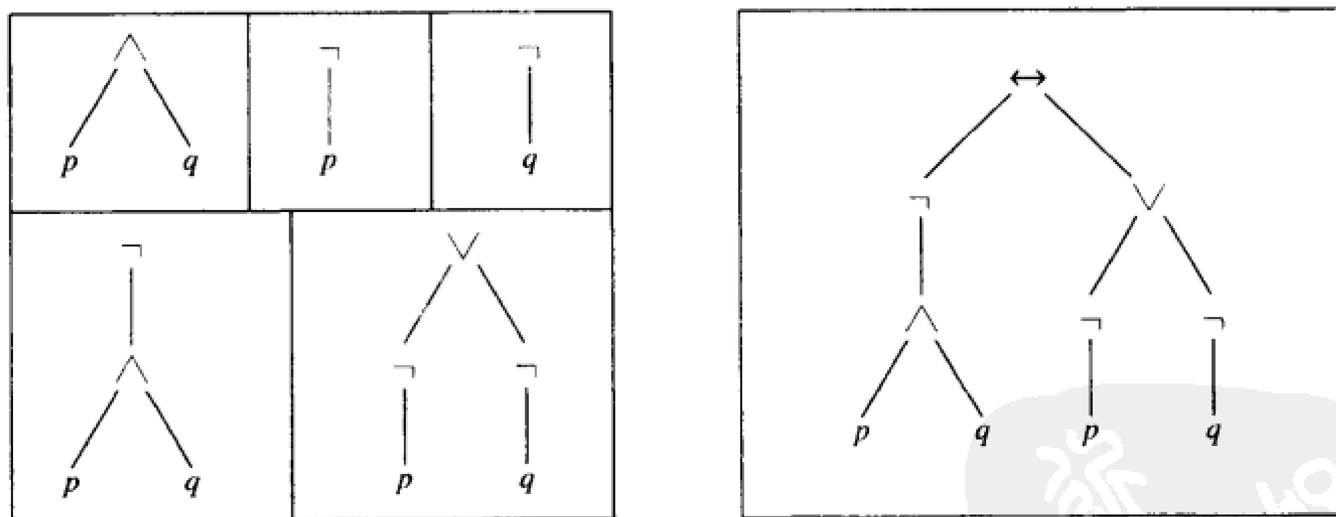


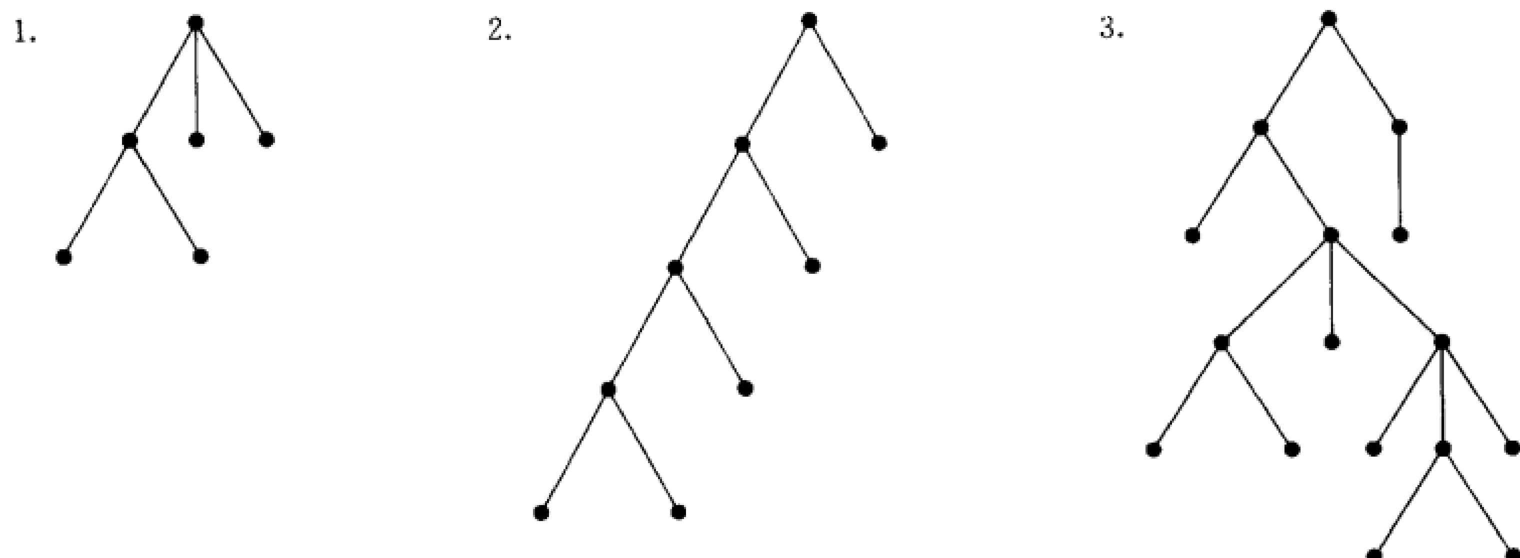
图 9-38 构造一个复合命题的根树

这个表达式的前缀、后缀和中缀形式是这样求出的：分别以前序、后序和中序来遍历这个根树(包含括号)。这些遍历分别给出$\leftrightarrow\neg\wedge pq\vee\neg p\neg q$，$pq\wedge\neg p\neg q\neg\vee\leftrightarrow$和$(\neg(p\wedge q))\leftrightarrow((\neg p)\vee(\neg q))$。 ■

因为前缀表达式和后缀表达式都是无二义性的，而且不用来回扫描就容易求出它们的值，所以它们在计算机科学里大量使用。这样的表达式对编译器的构造是特别有用的。

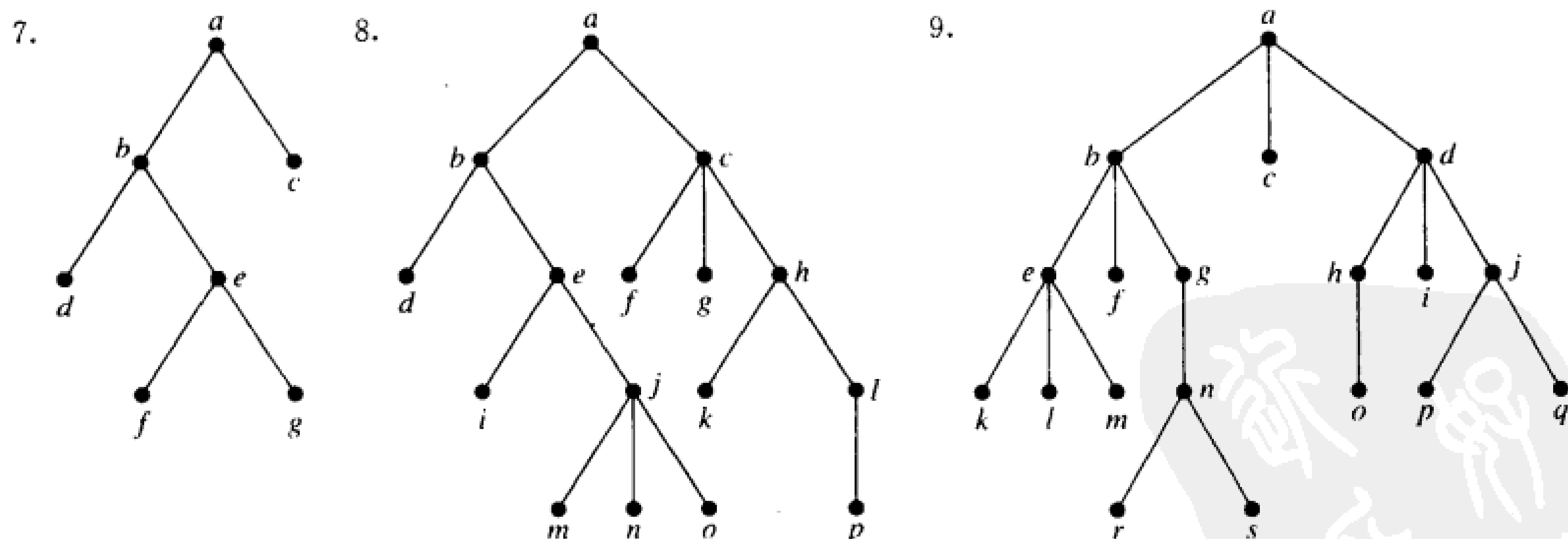
练习

在练习 1～3 里，对给定的有序根树构造通用地址系统。然后利用这个通用地址系统并且用顶点的标记的字典顺序来排序顶点。



4. 假定在有序根树 T 里顶点 v 的地址是 3.4.5.2.4。
   a) v 是在哪一层？
   b) v 的父母的地址是什么？
   c) v 的兄弟最少有多少？
   d) 若 v 具有这个地址，那么在 T 里最少可能有多少个顶点？
   e) 求其他必然出现的地址。
5. 假定在有序根树 T 里，地址最大的顶点的地址是 2.3.4.3.1。是否有可能确定 T 里的顶点数？
6. 有序根树的树叶能否具有下面的通用地址表？若能，则构造出这样的有序根树。
   a) 1.1.1，1.1.2，1.2，2.1.1.1，2.1.2，2.1.3，2.2，3.1.1，3.1.2.1，3.1.2.2，3.2
   b) 1.1，1.2.1，1.2.2，1.2.3，2.1，2.2.1，2.3.1，2.3.2，2.4.2.1，2.4.2.2，3.1，3.2.1，3.2.2
   c) 1.1，1.2.1，1.2.2，1.2.2.1，1.3，1.4，2，3.1，3.2，4.1.1.1

在练习 7～9 里，确定前序遍历访问所给的有序根树的顶点的顺序。



10. 使用中序遍历，以什么顺序访问练习 7 里有序根树的顶点？
11. 使用中序遍历，以什么顺序访问练习 8 里有序根树的顶点？
12. 使用中序遍历，以什么顺序访问练习 9 里有序根树的顶点？
13. 使用后序遍历，以什么顺序访问练习 7 里有序根树的顶点？
14. 使用后序遍历，以什么顺序访问练习 8 里有序根树的顶点？
15. 使用后序遍历，以什么顺序访问练习 9 里有序根树的顶点？

16. 用二叉树来表示表达式$((x+2)\uparrow 3)*(y-(3+x))-5$。表示方式应采用：
a)前缀记法。 b)后缀记法。 c)中缀记法。

17. 用二叉树来表示表达式$(x+xy)+(x/y)$和$x+((xy+x)/y)$。表示方式应采用：
a)前缀记法。 b)后缀记法。 c)中缀记法。

18. 用有序根树来表示复合命题$\neg(p\wedge q)\leftrightarrow(\neg p\vee\neg q)$和$(\neg p\wedge(q\leftrightarrow\neg p))\vee\neg q$。表示方式应采用：
a)前缀记法。 b)后缀记法。 c)中缀记法。

19. 用有序根树来表示$(A\cap B)-(A\cup(B-A))$。表示方式应采用：
a)前缀记法。 b)后缀记法。 c)中缀记法。

*20. 有多少种方式给字符串$\neg p\wedge q\leftrightarrow\neg p\vee\neg q$完全加上括号以便产生中缀表达式？

*21. 有多少种方式给字符串$A\cap B-A\cup B-A$完全加上括号以便产生中缀表达式？

22. 画出下面用前缀记法写出的每个算术表达式所对应的有序根树。然后用中缀记法来写每个表达式。
a) $+ * + - 5\ 3\ 2\ 1\ 4$ b) $\uparrow + 2\ 3 - 5\ 1$ c) $*\ /\ 9\ 3 + *\ 2\ 4 - 7\ 6$

23. 下面每个前缀表达式的值是什么？
a) $- *\ 2\ /\ 8\ 4\ 3$ b) $\uparrow - *\ 3\ 3 *\ 4\ 2\ 5$
c) $+ - \uparrow 3\ 2 \uparrow 2\ 3\ /\ 6 - 4\ 2$ d) $* + 3 + 3 \uparrow 3 + 3\ 3\ 3$

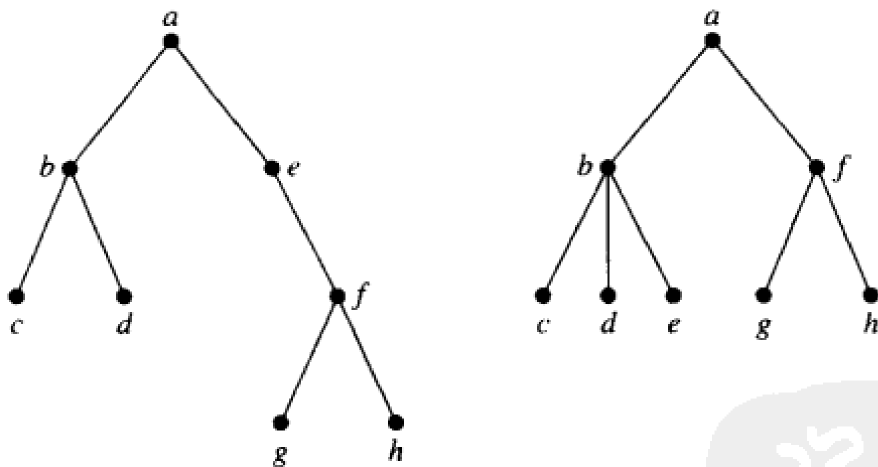
24. 下面每个后缀表达式的值是什么？
a) $5\ 2\ 1 - - 3\ 1\ 4 + + *$ b) $9\ 3\ /\ 5 + 7\ 2 - *$ c) $3\ 2 *\ 2 \uparrow 5\ 3 - 8\ 4\ / * -$

25. 构造前序遍历为$a, b, f, c, g, h, i, d, e, j, k, l$的有序根树，其中$a$有四个子女，$c$有三个子女，$j$有两个子女，$b$和$e$都有一个子女，所有其他顶点都是树叶。

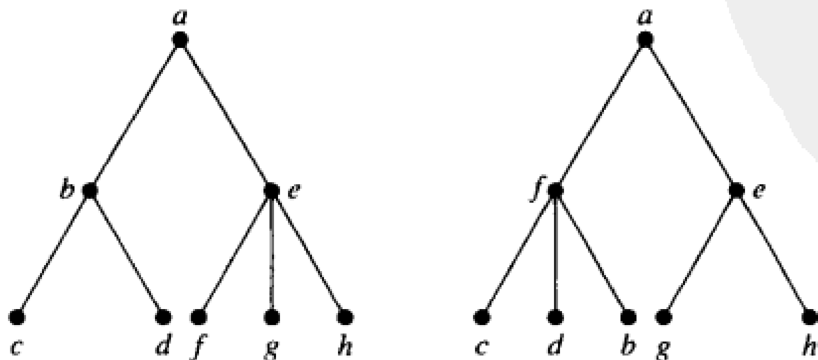
*26. 证明：当规定了有序根树的前序遍历所生成的顶点列表、并且规定了每个顶点的子女数时，这个有序根树是唯一确定的。

*27. 证明：当规定了有序根树的后序遍历所生成的顶点列表、并且规定了每个顶点的子女数时，这个有序根树是唯一确定的。

28. 证明：下图所示的两个有序根树的前序遍历产生相同的顶点列表。注意这个结果不与练习 26 里的命题相矛盾，因为在这两个有序根树里内点的子女数是不同的。



29. 证明：下图所示的两个有序根树的后序遍历产生相同的顶点列表。注意这个结果不与练习 27 里的命题相矛盾，因为在这两个有序根树里内点的子女数是不同的。



在一组符号和一组二元运算上用前缀记法表示的合式公式是用下面的规则来递归地定义的：

(i)若 x 是符号，则 x 是用前缀记法表示的合式公式；

(ii)若 X 和 Y 都是合式公式并且 $*$ 是运算，则 $*XY$ 是合式公式。

30. 下列哪些公式是在符号 $\{x, y, z\}$ 和二元运算集 $\{\times, +, \circ\}$ 上的合式公式？

a) $\times + + x\ y\ z$ b) $\circ\ x\ y \times x\ z$

c) $\times \circ\ x\ z \times \times\ x\ y$ d) $\times + \circ\ x\ x \circ\ x\ x\ x$

*31. 证明：在一组符号和一组二元运算上用前缀记法表示的任何合式公式所包含的符号数都比运算数恰好多一个。

32. 给出在一组符号和一组二元运算上用后缀记法表示的合式公式的定义。

33. 给出在符号 $\{x, y, z\}$ 和二元运算集 $\{+, \times, \circ\}$ 上带三个以上运算的、用后缀记法表示的合式公式的 6 个例子。

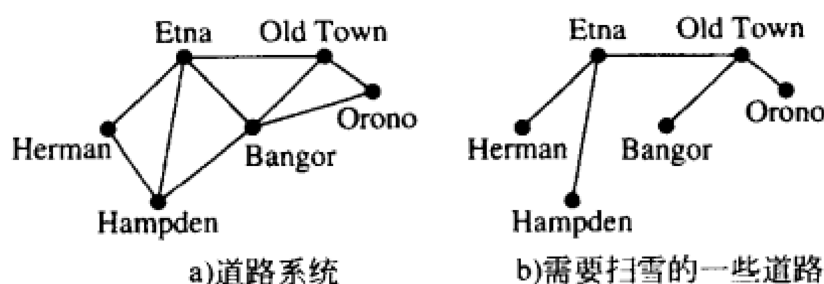
34. 把用前缀记法表示的合式公式的定义推广到这样的符号和运算集合上，其中运算可能不是二元的。

9.4 生成树

9.4.1 引言

考虑一下图 9-39a 所示的简单图所表示的缅因州的道路系统。在冬天里保持道路通畅的唯一方式就是经常扫雪。高速公路部门希望只扫尽可能少的道路上的雪，而确保总是存在连接任何两个乡镇的干净道路。如何才能做到这一点？

至少扫除 5 条道路上的雪才能保证在任何两个乡镇之间有一条通路。图 9-39b 说明了这样一些道路。注意表示这些道路的子图是树，因为它是连通的并且包含 6 个顶点和 5 条边。



a)道路系统 b)需要扫雪的一些道路

图 9-39

这个问题是用包含原来简单图的所有顶点、边数最小的连通子图来解决的。这样的图必然是树。

定义 1 设 G 是简单图。G 的生成树是包含 G 的每个顶点的 G 的子图。

有生成树的简单图必然是连通的，因为在任何两个顶点之间都有生成树里的通路。反过来也是对的；即每个连通图都有生成树。在证明这个结果之前将给出一个例子。

例 1 找出图 9-40 所示简单图的生成树。

解 图 G 是连通的，但它不是树，因为它包含简单回路。删除边 $\{a, e\}$。这样就消除了一个简单回路，而且所得出的子图仍然是连通的并且仍然包含 G 的每个顶点。其次删除边 $\{e, f\}$ 以便消除第二个简单回路。最后，删除边 $\{c, g\}$ 以便产生一个没有简单回路的简单图。这个子图是生成树，因为它是包含 G 的每个顶点的树。图 9-41 中说明了用来产生这个生成树的边的删除序列。

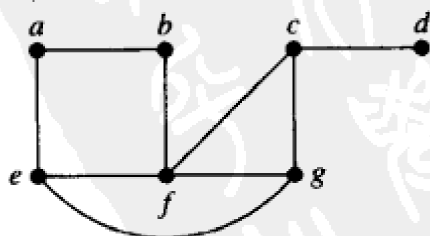


图 9-40 简单图 G

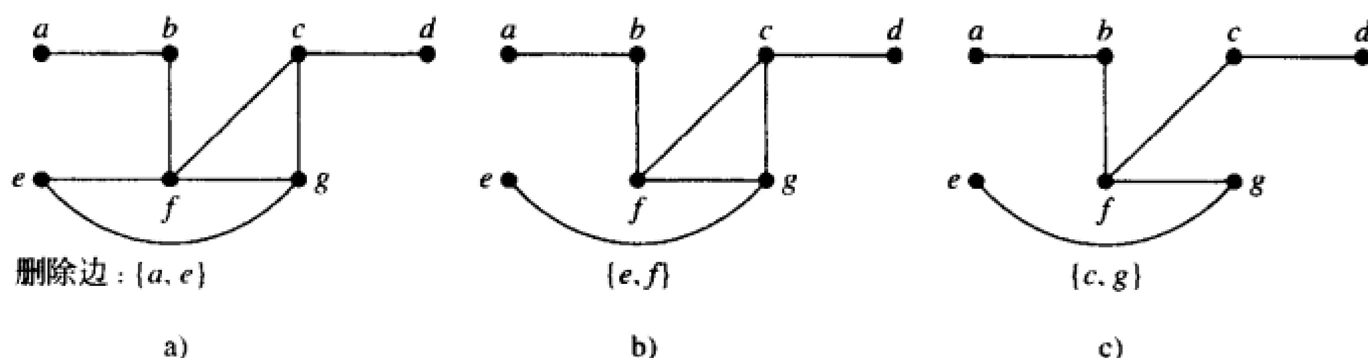


图 9-41　通过删除形成简单回路的边来产生 G 的一个生成树

图 9-41 所示的生成树不是唯一的 G 的生成树。例如，图 9-42 所示的每个树都是 G 的生成树。■

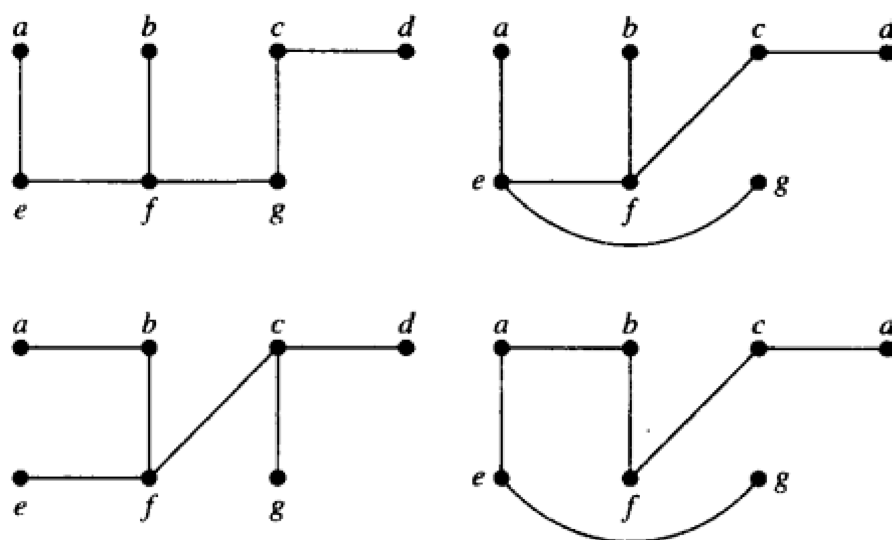


图 9-42　G 的一些生成树

定理 1　简单图是连通的当且仅当它具有生成树。

证　首先，假定简单图 G 有生成树 T。T 包含 G 的每个顶点。另外，在 T 的任何两个顶点之间都有 T 里的通路。因为 T 是 G 的子图，所以在 G 的任何两个顶点之间都有 T 里的通路。因此，G 是连通的。

现在假定 G 是连通的。若 G 不是树，则它必然包含简单回路。从这些简单回路中的一个里删除一条边。所得出的子图少了一条边，但是仍然包含 G 的所有顶点并且是连通的。若这个子图不是树，则它有简单回路；所以像前面那样，删除一个简单回路里的一条边。重复这个过程直到没有简单回路为止。这是可能的，因为在图里只有有穷的边数。当没有简单回路剩下时，这个过程终止。产生出一个树，因为在删除边时这个图保持连通。这个树是生成树，因为它包含 G 的每个顶点。□

例 2 说明，在数据网络里生成树是重要的。

例 2　**IP 组播**　在网际协议(IP)网络上的组播里，生成树的作用很重要。为了从源计算机发送数据到多个接收计算机，每个接收计算机是一个子网，可以分别发送数据到每个计算机。称为单点广播的这种类型的网络是无效的，这是因为在网络上存有发送相同数据的多个副本。为了更有效地传送数据到多个接收计算机，就使用 IP 组播。在 IP 组播里，一个计算机在网络上发送数据的单一副本，当数据到达中间路由器时，就把数据分发到一个或更多的其他路由器，以便接收计算机都在它们不同的子网里最终接收到这些数据。(路由器是专门在网络里子网之间分发 IP 数据报文的计算机。在组播里，路由器使用 D 类地址，每个都表示接收计算机可以加入的一个会话；见 4.1 节例 15。)

为了让数据尽可能快地到达接收计算机，在数据穿过网络的通路里就不应当存在环路(在图论名词里它们是回路)。即一旦数据已经到达一个具体的路由器，数据就再也不应当返回这个路由器。为了避免环路，组播路由器用网络算法来构造图 9-43 所示的生成树，这个图以组播源、路由器和包含接收计算机的子网来作为顶点，以边表示计算机和(或)路由器之间的连接。这个生成树的根就是组播源。包含接收计算机的子网就是这个树的树叶(注意不包含接收计算机的子网都不包含在这个图里)。在图 9-43 里说明这些内容。 ■

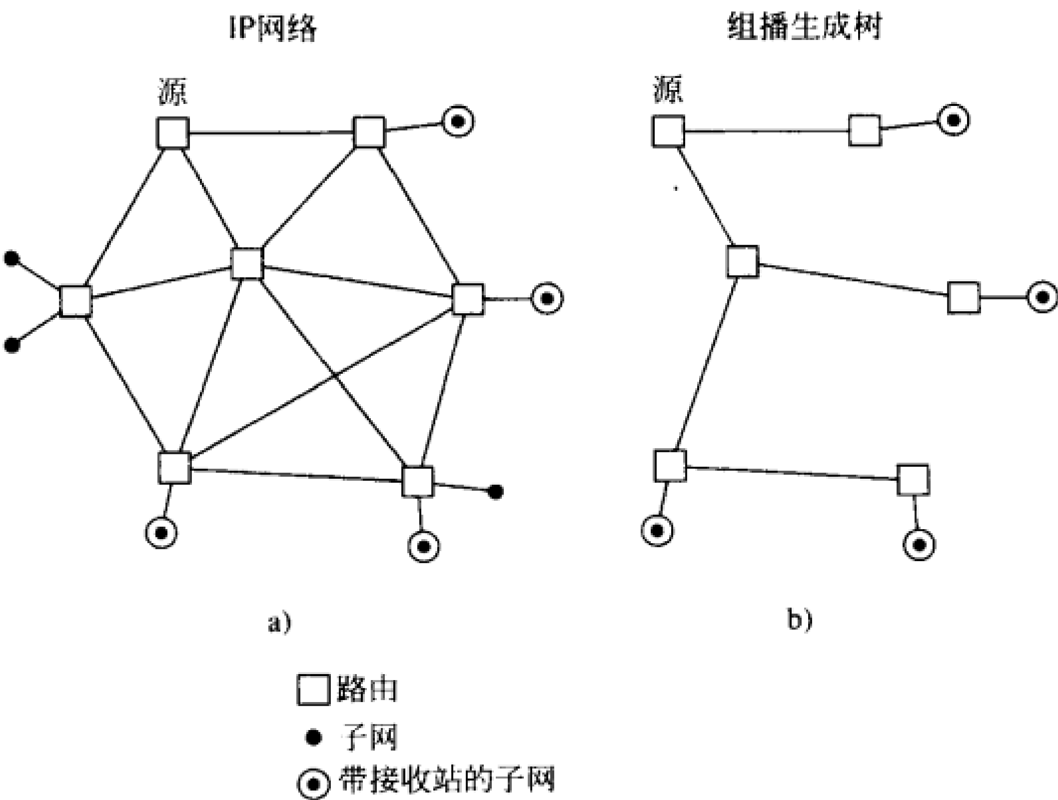


图 9-43　一个组播生成树

9.4.2 深度优先搜索

定理 1 的证明给出通过从简单回路删除边来找出生成树的算法。这个算法是低效的，因为它要求找出简单回路。另一种不采用删除边来构造生成树的方法是，通过相继地添加边来建立生成树。在这里将给出基于这个原理的两个算法。

可以用深度优先搜索来建立连通简单图的生成树。我们将形成一个根树，而生成树将是这个根树的无向底图。任意选择图中一个顶点作为根。通过相继地添加边来形成从这个顶点开始的通路，其中每条新边都与通路上的最后一个顶点以及还不在通路上的一个顶点相关联。继续尽可能地添加边到这条通路。若这条通路经过图的所有顶点，则由这条通路组成的树就是生成树。不过，若这条通路没有经过图的所有顶点，则必须添加其他的边。退到通路里的倒数第二个顶点，若有可能，则形成从这个顶点上开始的经过还没有访问过的顶点的通路。若不能这样做，则后退到通路里的另外一个顶点，即在通路里后退两个顶点，然后再试。

重复这个过程，从所访问过的最后一个顶点开始，在通路上一次后退一个顶点，只要有可能就形成新的通路，直到不能添加更多的边为止。因为这个图有有穷的边数并且是连通的，所以这个过程以产生生成树而告终。在这个算法的一个阶段上是通路末端的顶点将是根树里的树叶，而在其上开始构造一条通路的顶点将是内点。

读者应当注意到这个过程的递归本质。另外，注意若图中的顶点是排序的，则当总是选择在该顺序里可用的第一个顶点时，在这个过程的每个阶段上对边的选择就全都是确定的。不过，将不总是明显地对图的顶点排序。

深度优先搜索也称为回溯，因为这个算法返回以前访问过的顶点以便添加边。下面的例子说明回溯。

例 3 用深度优先搜索来找出图 9-44 所示图 G 的生成树。

Extra Examples **解** 在图 9-45 里显示深度优先搜索为了产生 G 的生成树而使用的步骤。任意地从顶点 f 开始。一条通路是这样建立的：相继地添加与还不在通路上的顶点相关联的边，只要有可能就这样做。这样就产生通路 f，g，h，k，j(注意也可能建立其他的通路)。下一步，回溯到 k。不存在从 k 开始包含还没有访问过的顶点的通路。所以回溯到 h。形成通路 h，i。然后回溯到 g，然后再回溯到 f。从 f 建立通路 f，d，e，c，a。然后再回溯到 c 并且形成通路 c，b。这样就产生了生成树。 ■

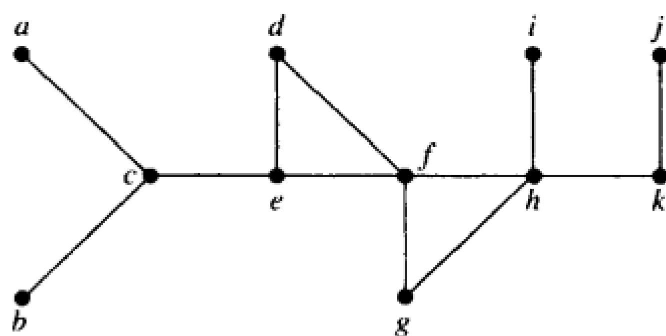


图 9-44 图 G

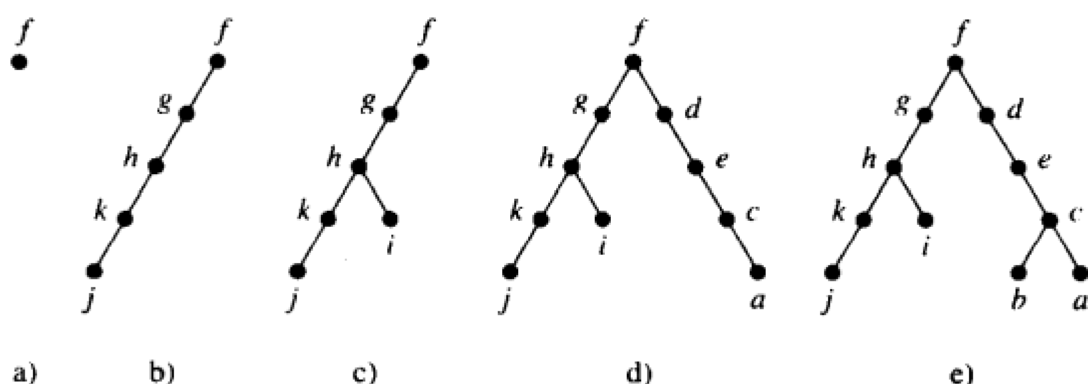


图 9-45 G 的深度优先搜索

一个图的深度优先搜索所选择的边都称为树边。这个图所有其他的边都必然连接一个顶点与这个顶点在树中的祖先或后代。这些边都称为背边(练习 39 要求证明这个事实)。

例 4 图 9-46 中强调了从顶点 f 开始的深度优先搜索所找到的树边，用粗线显示这些树边。用细黑线显示背边(e，f)和(f，h)。 ■

我们已经解释了如何用深度优先搜索来求图的生成树。但是，迄今为止的讨论还没有带出深度优先搜索的递归本质。为了弄清深度优先搜索的递归本质，需要几个术语。当执行深度优先搜索的步骤时，当把顶点 v 加入树时说从顶点 v 开始探索，当最后一次回溯回到 v 时说从顶点 v 结束探索。理解算法的递归本质所需要的关键事实是，当加入连接顶点 v 到顶点 w 的边时，在回到 v 完成从 v 的探索之前就结束了从 w 的探索。在算法 1 中构造了带顶点 v_1，…，v_n 的图 G 的生成树，首先选择顶点 v_1 作为树根。开始时令 T 是只有这一个顶点的树。在每个步骤，加入一个新顶点到 T 以及从已在 T 中的一个顶点到这个新顶点的一条边，并且从这个新顶点探索。注意当算法完成时，T 没有简单回路，因为没有加入连接到已在树中的顶点的边。另外，T 在构造时保持连通(用数学归纳法可以轻而易举地证明最后这两个事实)。由于 G 是连通的，G 的每个顶点都被算法访问到并且加入到树中(读者可以验证)。所以 T 是 G 的生成树。

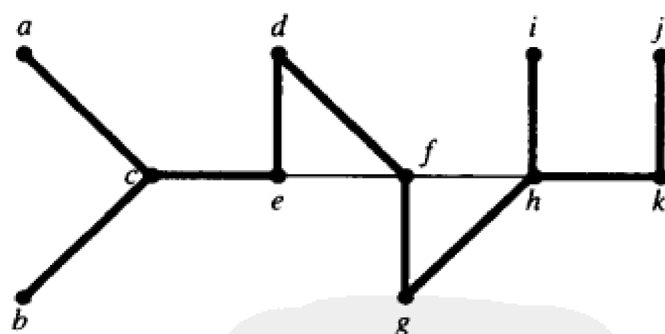


图 9-46 例 4 中深度优先搜索的树边和背边

```
算法 1 深度优先搜索
procedure DFS(G: 带顶点 v1, …, vn 的连通图)
T := 只包含顶点 v1 的树
visit(v1)
procedure visit(v: G 的顶点)
for 与 v 相邻并且还不在 T 中的每个顶点 w
begin
  加入顶点 w 和边{v, w}到 T
  visit(w)
end
```

现在分析深度优先搜索算法的计算复杂性。关键事实是对于每个顶点 v 来说，当在搜索中首次遇到顶点 v 时，就调用过程 $visit(v)$ 并且以后不再调用这个过程。假设 G 的邻接表是可用的(参见 8.3 节)，求出与 v 相邻的顶点不需要任何计算。当遵循算法的步骤时，至多检查每条边两次以确定是否加入这条边及其一个端点到树中。因此，过程 DFS 用 $O(e)$ 或 $O(n^2)$ 个步骤来构造一个生成树，其中 e 和 n 分别是 G 的边数和顶点数。(注意一个步骤包括：检查一个顶点是否已在正在构造的树中，如果这个顶点还不在树中，则加入这个顶点和对应的边。还利用了不等式 $e \leqslant n(n-1)/2$，对于任意简单图来说这个不等式都成立。)

深度优先搜索可以作为解决许多不同问题的算法的基础。例如，可以用来求图的连通分支并且可以用来求连通图的割点。将要看到，深度优先搜索是用来搜索计算困难问题的解的回溯技术的基础(参见[GrYe99]、[Ma89]和[CoLeRiSt01]对基于深度优先搜索的算法的讨论)。

9.4.3 宽度优先搜索

Demo 也可以通过使用宽度优先搜索来产生简单图的生成树。同样，将构造一个根树，而这个根树的无向底图就形成生成树。从图的顶点中任意地选择一个根。然后添加与这个顶点相关联的所有边。在这个阶段所添加的新顶点成为生成树在 1 层上的顶点。将新顶点任意排序。下一步，按顺序访问 1 层上的每个顶点，只要不产生简单回路，就将与这个顶点相关联的每条边添加到树里。这样就产生了树在 2 层上的顶点。遵循相同的过程，直到已经添加了树里的所有顶点。因为在图中的边数是有限的，所以这个过程会终止。在产生了包含图中每一个顶点的树之后，生成树也就产生了。下面是宽度优先搜索的一个例子。

例 5 用宽度优先搜索来找出图 9-47 所示的图的生成树。

Extra Examples **解** 在图 9-48 里显示宽度优先搜索过程的各步骤。选择顶点 e 作为根。然后添加与 e 相关联的所有边，所以添加了从 e 到 b，d，f 和 i 的边。这四个顶点都是在树的 1 层上。下一步，添加从 1 层上的顶点到还不在树上的相邻顶点的边。因此，添加从 b 到 a 和 c 的边，以及从 d 到 h，从 f 到 j 和 g，从 i 到 k 的边。新顶点 a，c，h，j，g 和 k 都是在 2 层上。下一步，添加从这些顶点到还不在树上的相邻顶点的边。这样就添加从 g 到 l 以及从 k 到 m 的边。■

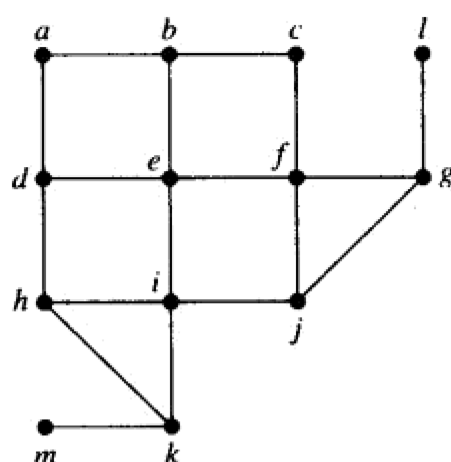


图 9-47 图 G

算法 2 的伪码描述了宽度优先搜索。在这个算法中假设连通图 G 的顶点排序为 v_1，…，v_n。在算法中，只要还没有产生简单回路就加入新的顶点和对应的边到树中，与正在处理的当前顶点相邻。

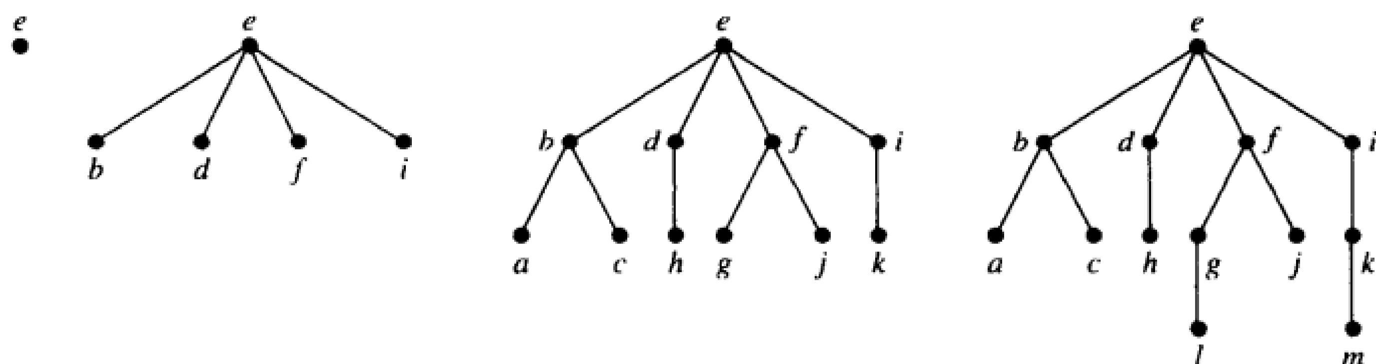


图 9-48　G 的宽度优先搜索

算法 2　宽度优先搜索

procedure BFS(G: 带顶点 v_1, …, v_n 的连通图)
T := 只包含顶点 v_1 的树
L := 空表
把 v_1 放入尚未处理顶点的表 L 中
while L 非空
begin
　删除 L 中第一个顶点 v
　for v 的每个邻居 w
　　if w 既不在 L 中也不在 T 中 **then**
　　begin
　　　加入 w 到表 L 的末尾
　　　加入 w 和边 $\{v, w\}$ 到 T
　　end
end

现在分析宽度优先搜索的计算复杂性。对于图中的每个顶点 v 来说，检查所有与 v 相邻的顶点并且加入每个尚未访问过的顶点到树 T 中。假设图的邻接表是可用的，确定哪些顶点与给定顶点相邻就不需要任何计算。如同在深度优先搜索算法的分析中那样，我们检查每条边至多两次来确定是否应当加入这条边及其尚未在树中的端点。所以宽度优先搜索算法使用 $O(e)$ 或 $O(n^2)$ 个步骤。

9.4.4　回溯

有些问题只能通过执行对所有可行解的穷举搜索来解决。系统地搜索出一个解的一种方式是使用决策树，其中每个内点都表示一次决策，而每个树叶都表示一个可行解。为了通过回溯来求出一个解，首先尽可能地做出一系列决策来尝试得出一个解。可以用决策树里的通路来表示决策序列。一旦知道了决策序列的任何扩展都不能得出解，就回溯到父母顶点并且若有可能，则用另外一个决策序列来尝试得出一个解。继续这个过程，直到找到一个解，或者证明没有解存在为止。下面的例子说明回溯的有用性。

例 6　**图着色**　如何用回溯来判定是否可以用 n 种颜色给一个图着色？

解　以下面的方式用回溯来解决这个问题。首先选择某个顶点 a 并且指定它颜色为 1。然后挑选第二个顶点 b，而且若 b 不与 a 相邻，则指定它颜色为 1。否则，指定 b 颜色为 2。然后来到第三个顶点 c。若有可能则对 c 用颜色 1。否则若有可能则用颜色 2。只有当颜色 1 和颜色 2 都不能用时才使用颜色 3。继续这个过程，只要有可能就为每个新顶点指定 n 种颜色中的一种，而且总是使用表中第一种允许的颜色。若遇到不能用 n 种颜色中任何一种来着色的顶点，则回溯到最后一次所做的指定，并且若有可能就改变最后着色的顶点的颜色，用表中下一种允许的颜色。若不可能改变这个颜色，则再回溯到更前面的指定，一次后退一步，直到有可能改变一个顶点的颜

色为止。然后只要有可能就继续指定新顶点的颜色。若使用 n 种颜色的着色存在，则可以通过回溯来产生(但是这个过程是极其低效的)。

具体地说，考虑用三种颜色来着色图 9-49 所示的图。图 9-49 所示的树说明如何用回溯来构造 3 着色。在这个过程里，首先用红色，其次用蓝色，最后用绿色。显然用回溯来解决这个简单的例子只是为了能够比较好地说明这项技术。

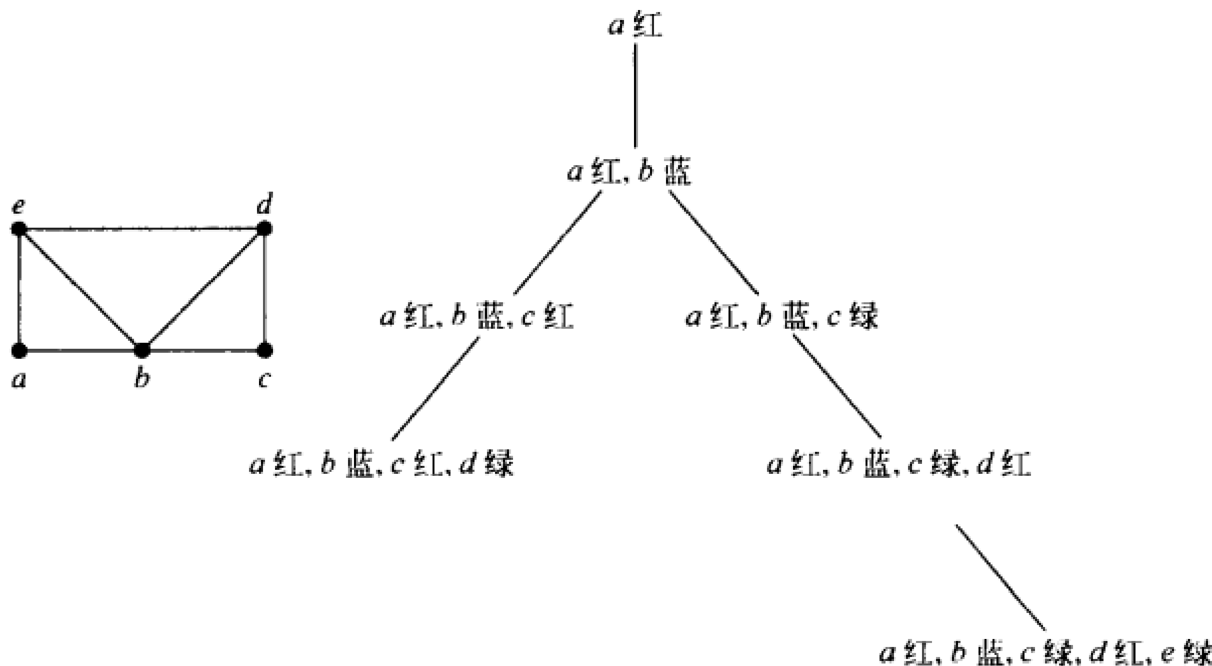


图 9-49 用回溯给图着色

在这个树里，从根开始的、表示指定红色给 a 的最初的通路，导致 a 红色、b 蓝色、c 红色而 d 绿色的着色。当以这种方式来着色 a，b，c 和 d 时，就不可能用这三种颜色中的任何一种来着色 e。所以，回溯到表示这个着色的顶点的父母。因为没有其他颜色可以用在 d 上，所以再回溯一层。然后改变 c 的颜色为绿色。通过接着指定红色给 d 和绿色给 e，就获得这个图的着色。■

例 7 ***n* 皇后问题** n 皇后问题问：在 $n \times n$ 棋盘上如何放置 n 个皇后，使得没有两个皇后可以互相攻击。如何用回溯来解决 n 皇后问题?

解 为了解决这个问题，必须在 $n \times n$ 棋盘上找出 n 个位置，使得这些位置中没有两个是在同一行上、同一列上或在同一斜线上(斜线是由对某个 m 来说满足 $i+j=m$ 或对某个 m 来说满足 $i-j=m$ 的所有位置 (i, j) 组成的)。将用回溯来解决 n 皇后问题。从空棋盘开始。在 $k+1$ 阶段，尝试在棋盘上第 $k+1$ 列里放置一个新皇后，其中在前 k 列里已经有了皇后。检查第 $k+1$ 列里的格子，从第一行的格子开始，寻找放置这个皇后的位置，使得它不与已经在棋盘上的皇后在同一行里或在同一斜线上(已经知道它不在同一列里)。若不可能在第 $k+1$ 列里找到放置皇后的位置，则回溯到在第 k 列里对皇后的放置。在这一列里下一个允许的行里放置皇后，若这样的行存在的话。若没有这样的行存在，则继续回溯。

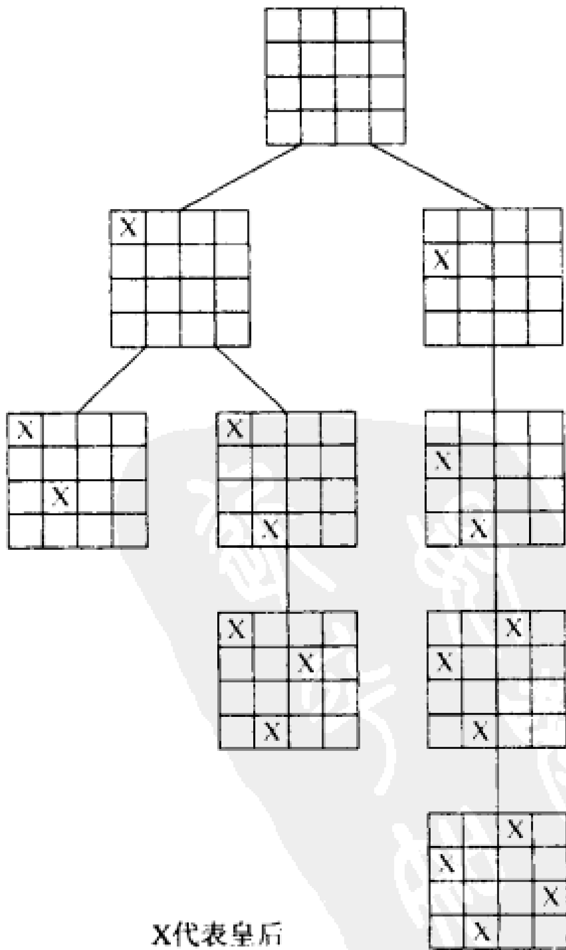


图 9-50 四皇后问题的回溯解法

具体地说，图 9-50 显示四皇后问题的回溯解法。在这种解法里，在第一行第一列里放置一个皇后。然后在第二列的第三行里放置一个皇后。不过，这样就使得不可能在

第三列里放置一个皇后。所以就回溯并且在第二列的第四行里放置一个皇后。当这样做时，就可以在第三列的第二行里放置一个皇后。但是没有办法在第四列里添加一个皇后。这说明当在第一行第一列里放置一个皇后时就得不出解。回溯到空棋盘，在第一列的第二行里放置一个皇后。这样就得出图 9-50 所示的解。■

例 8　子集和　考虑下面的问题。给定一组正整数 x_1，x_2，…，x_n，求这组整数的和为 M 的一个子集。如何用回溯来解决这个问题？

解　从空无一项的和来开始。通过相继地添加项来构造这个和。若当添加这个序列里的一个整数到和里而这个和仍然小于 M 时，则包含这个整数。若得出使得添加任何一项就大于 M 的一个和，则通过去掉这个和的最后一项来回溯。

图 9-51 显示下面这个问题的回溯解法，求{31，27，15，11，7，5}的和等于 39 的子集。■

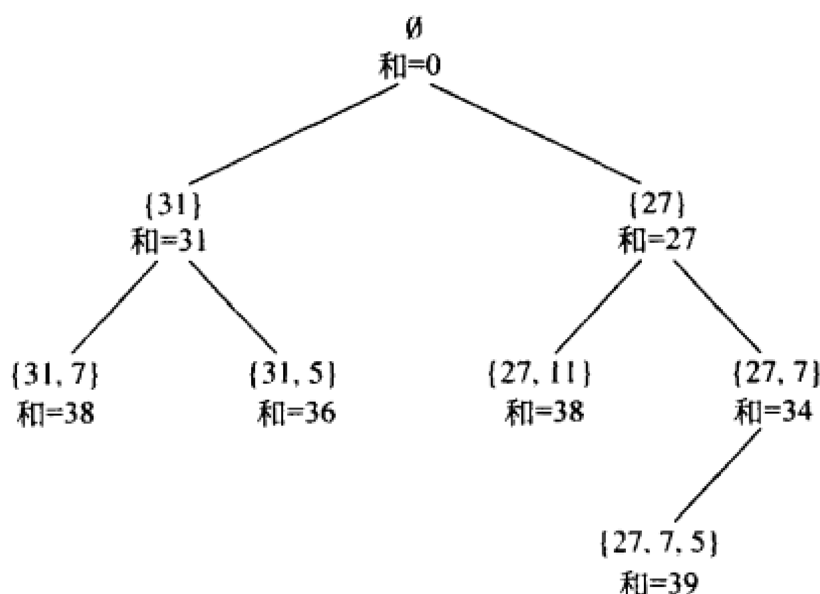


图 9-51　用回溯求等于 39 的和

9.4.5　有向图中的深度优先搜索

可以轻而易举地修改深度优先搜索和宽度优先搜索，使得以有向图作为输入时它们也能运行。但是，输出不一定是生成树，而可能是森林。在这两个算法中，只有当一条边从正在考虑的顶点发出并且到一个尚未加入的顶点时才加入这条边。如果在其中任何一个算法的某个阶段找不到从已经加入的顶点到尚未加入的顶点的边，则算法加入的下一个顶点成为生成森林中一个新树的根。这一点在例 9 中解释。

例 9　给定图 9-52a 所示的图作为输入，深度优先搜索的输出是什么？

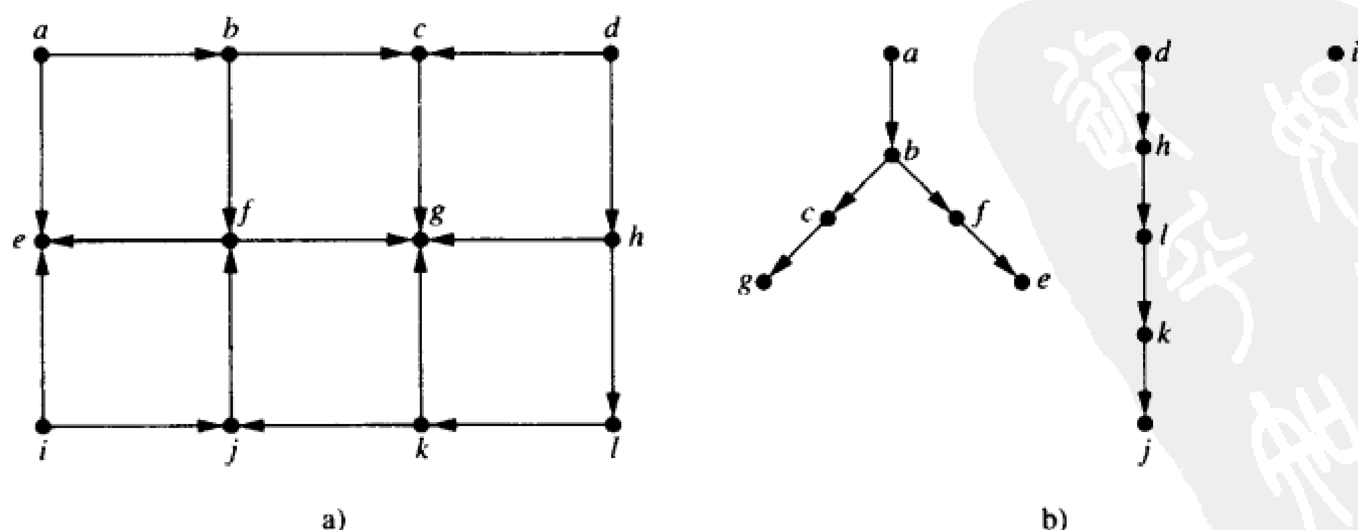


图 9-52　有向图的深度优先搜索

解 从顶点 a 开始深度优先搜索并且加入顶点 b、c 和 g 以及无路可走时对应的边。回溯到 c，但是仍然无路可走，于是回溯到 b，在这里加入顶点 f 和 e 以及对应的边。回溯最终又回到 a。然后在 d 开始一个新的树并且加入顶点 h、l、k 和 j 以及对应的边。回溯到 k，然后到 l，然后到 h 并且回到 d。最后，在 i 开始一个新的树，完成深度优先搜索。输出如图 9-52b 所示。■

有向图中的深度优先搜索是许多算法的基础(参见[GrYe99]、[Ma89]和[CoLeRiSt01])。它可以用来确定有向图是否具有回路，可以用来完成图的拓扑排序，也可以用来求有向图的强连通分支。

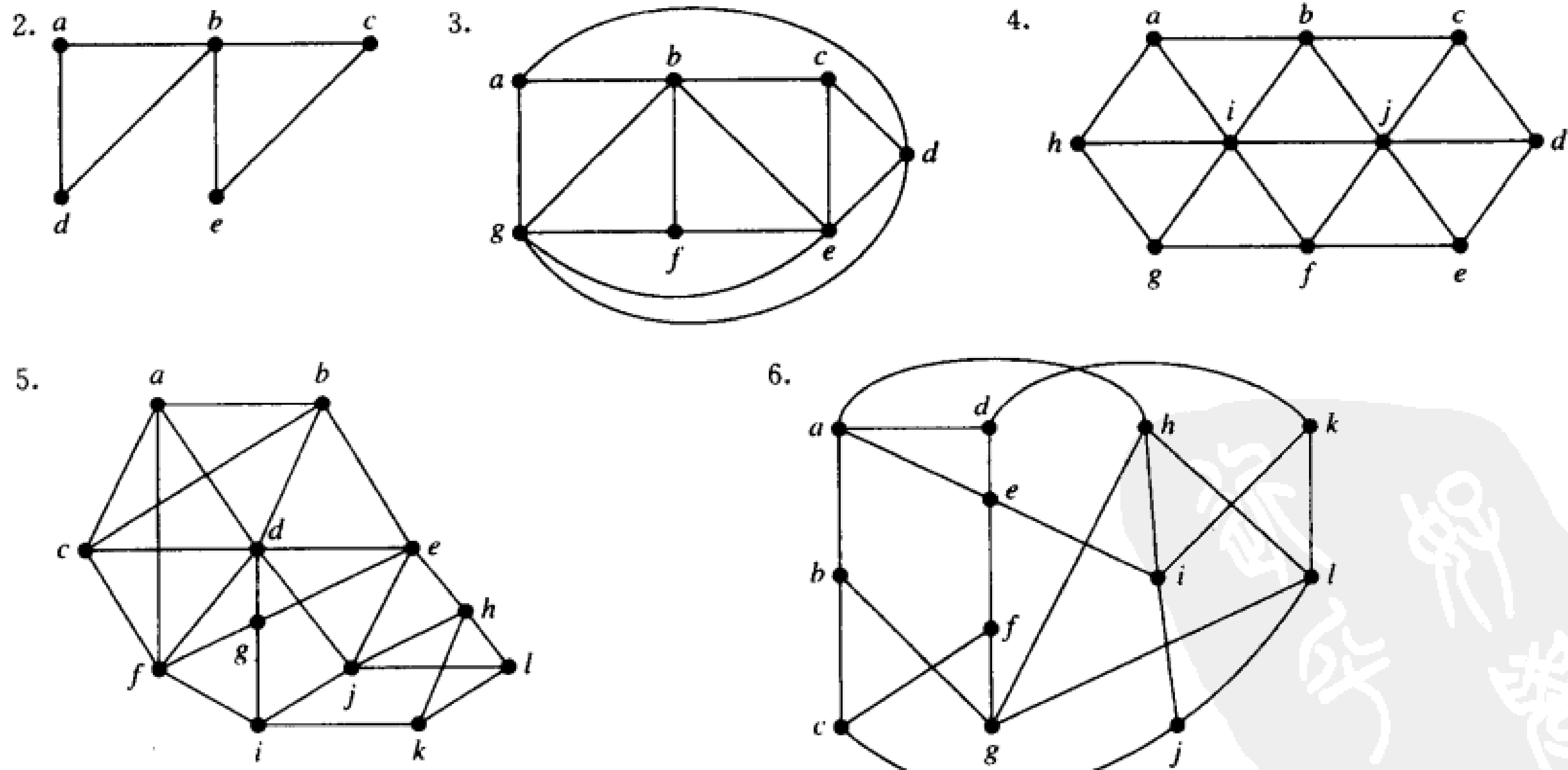
用深度优先搜索和宽度优先搜索在网络搜索引擎上的应用来结束本节。

例 10 网络蜘蛛 为了给网站建立索引，诸如 Google、Hotbot 和 Lycos 这些著名的搜索引擎从已知的站点开始系统地探索网络。这些搜索引擎使用所谓网络蜘蛛(或网络爬虫、网络机器人)的程序来访问网站并且分析其内容。网络蜘蛛同时使用深度优先搜索和宽度优先搜索来创建索引。如 8.1 节例 8 所述，可以用所谓网络图的有向图来为网页和网页之间的链接建立模型。用顶点表示网页并且用有向边表示链接。利用深度优先搜索，选择一个初始的网页，循着一个链接(如果存在这样的链接的话)到达第二个网页，循着第二个网页的一个链接(如果存在这样的链接的话)到达第三个网页，等等，直到找到一个没有新的链接的网页为止。然后使用回溯来检查前面阶段的链接去寻找新的链接，等等。(由于实际限制，网络蜘蛛在深度优先搜索中的搜索深度是有限的。)利用宽度优先搜索，选择一个初始的网页并且遵循这个网页上的一个链接到达第二个网页，然后遵循初始网页上的第二个链接(如果存在的话)，等等，直到已经遵循了初始网页上的所有链接为止。然后逐页地遵循下一层网页上的链接，等等。■

练习

1. 为了产生生成树，必须从带有 n 个顶点和 m 条边的连通图里删除多少条边？

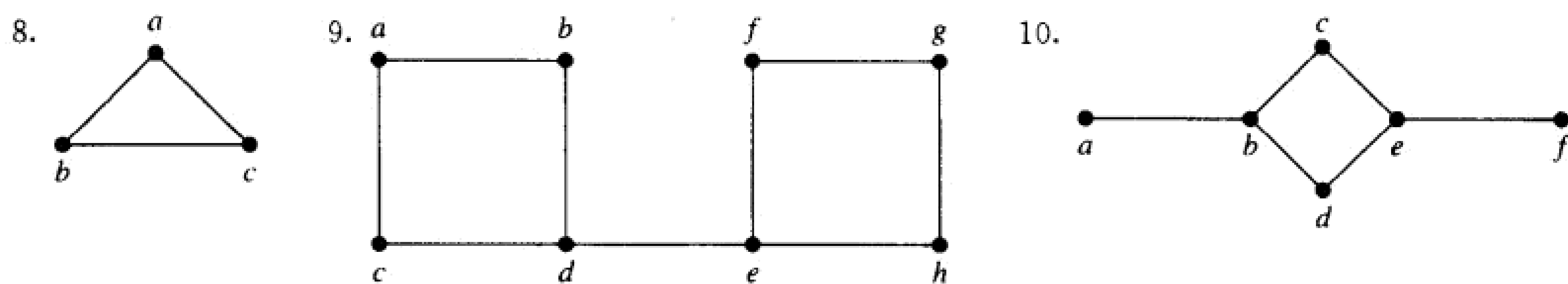
在练习 2～6 里，通过删除简单回路里的边来求所示的图的生成树。



7. 求下面每个图的生成树。

a) K_5　b) $K_{4,4}$　c) $K_{1,6}$　d) Q_3　e) C_5　f) W_5

在练习 8～10 里，画出所给的简单图的生成树。



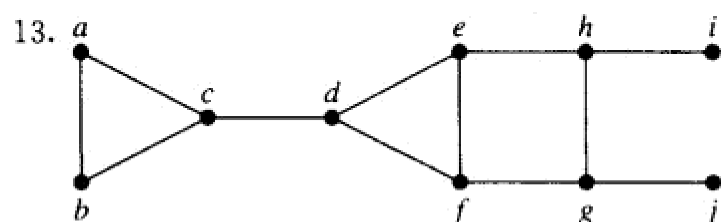
*11. 下面的每个简单图各有多少个不同的生成树？

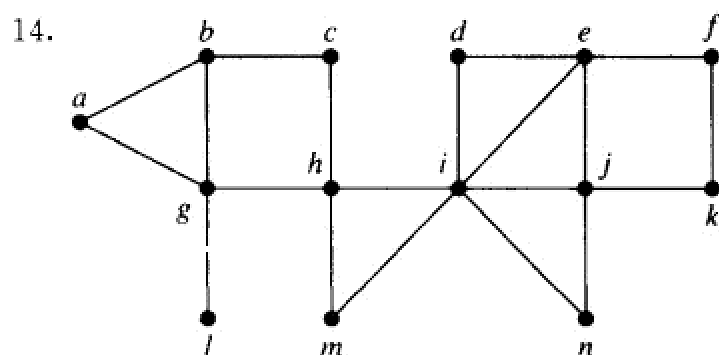
a) K_3　　b) K_4　　c) $K_{2,2}$　　d) C_5

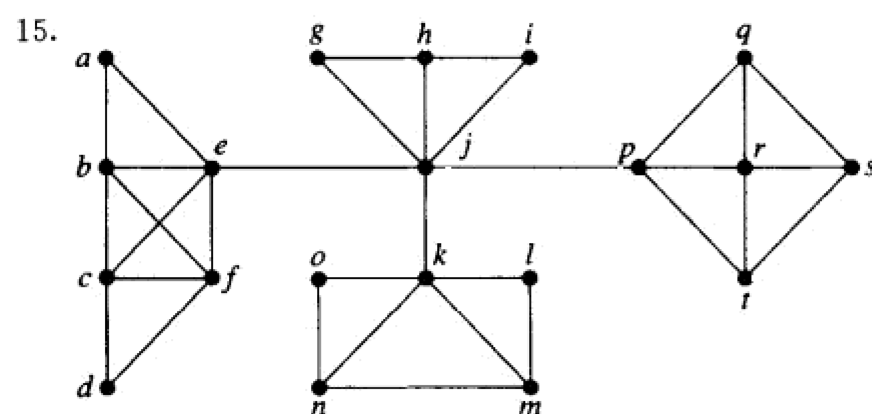
*12. 下面的每个简单图各有多少种不同构的生成树？

a) K_3　　b) K_4　　c) K_5

在练习 13～15 里，用深度优先搜索来构造所给的简单图的生成树。选择 a 作为这个生成树的根并且假定顶点都以字母顺序来排序。







16. 用宽度优先搜索来构造练习 13～15 里每个简单图的生成树。选择 a 作为每个生成树的根。

17. 用深度优先搜索求下列这些图的生成树。

a) W_6（参见 8.2 节例 6），从 6 度顶点开始　　b) K_5

c) $K_{3,4}$，从 3 度顶点开始　　d) Q_3

18. 用宽度优先搜索求练习 17 中的每个图的生成树。

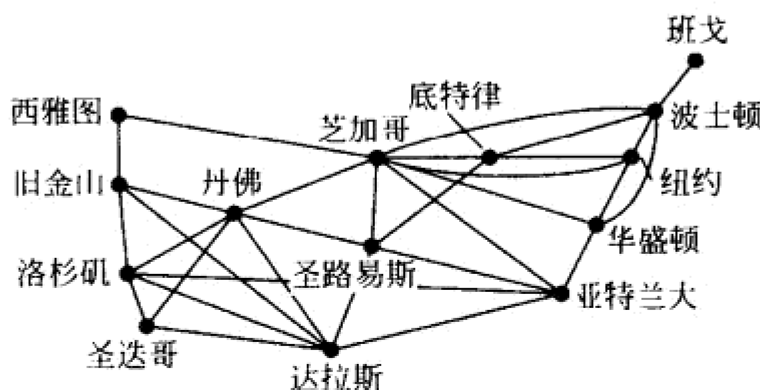
19. 描述一下轮图 W_n 的宽度优先搜索和深度优先搜索所产生的树，从 n 度顶点开始，其中 n 是整数满足 $n \geqslant 3$（参见 8.2 节例 6）。说明答案的合理性。

20. 描述一下完全图 K_n 的宽度优先搜索和深度优先搜索所产生的树，从 m 度顶点开始，其中 n 是正整数。说明答案的合理性。

21. 描述一下完全偶图 $K_{m,n}$ 的宽度优先搜索和深度优先搜索所产生的树，其中 m 和 n 都是正整数。说明答案的合理性。

22. 解释如何用宽度优先搜索和深度优先搜索来确定一个图是否为偶图。

23. 假定一家航空公司必须压缩它的航班以节省资金。若它原来的航线如下图所示，则可以中断哪些飞行以保持在所有城市对之间的服务（其中从一个城市飞往另一个城市可能需要换乘飞机）？



24. 何时连通简单图里的一条边必然在这个图的每个生成树里？
25. 哪些连通简单图恰好有一个生成树？
26. 解释一下如何用宽度优先搜索或深度优先搜索来排序连通图的顶点。
*27. 证明：在连通简单图里顶点 v 和 u 之间的最短通路的长度，等于在以 v 为根的 G 的宽度优先生成树里 u 的层数。
28. 用回溯来试验使用三种颜色对 8.8 节练习 5～7 里每个图的着色。
29. 用回溯来对下面的 n 值解决 n 皇后问题。

a) $n=3$　　b) $n=5$　　c) $n=6$

30. 用回溯来求集合{27，24，19，14，11，8}的和为下列值的子集，若存在的话。

a) 20　　b) 41　　c) 60

31. 解释一下如何用回溯来找出图中的哈密顿通路或哈密顿回路。
32. a) 解释一下如何用回溯来找出迷宫的出路，给定出发位置和出口位置。考虑把迷宫划分成位置，其中在每个位置上的移动包括四种可能性之一（上，下，右，左）。

b) 找出在下面的迷宫里从标记为 X 的出发位置到出口的通路。

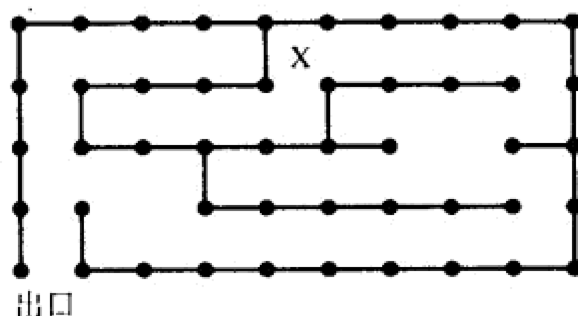


图 G 的生成森林是包含 G 的每个顶点的森林，使得当在两个顶点之间有 G 里的通路时，这两个顶点就在同一个树里。

33. 证明：每个有穷简单图都有生成森林。
34. 在图的生成森林里有多少个树？
35. 对带有 n 个顶点、m 条边和 c 个连通分支的图来说，必须删除多少条边才能产生它的生成森林？
36. 设计基于删除形成简单回路的边来构造图的生成森林的算法。
37. 设计基于深度优先搜索来构造图的生成森林的算法。
38. 设计基于宽度优先搜索来构造图的生成森林的算法。
39. 设 G 是连通图。证明：如果 T 是用深度优先搜索构造的 G 的生成树，则 G 的不在 T 中的边必定是背边，换句话说，这条边必定连接一个顶点到这个顶点在 T 中的祖先或后代。
40. 设 G 是连通图。证明：如果 T 是用宽度优先搜索构造的 G 的生成树，则 G 的不在 T 中的边必定连接这个生成树中在同一层上或相差一层的顶点。
41. 对于哪些图来说，无论选择哪个顶点作为树根，深度优先搜索和宽度优先搜索都产生同样的生成树？说明答案的合理性。
42. 用练习 39 证明：如果 G 是含有 n 个顶点的连通简单图并且 G 不含长度为 k 的简单通路，则 G 至多含有 $(k-1)m$ 条边。

43. 用数学归纳法证明：宽度优先搜索按照顶点在所得出的生成树中的层数的顺序来访问这些顶点。
44. 用伪码来描述深度优先搜索的一个变种，它把整数 n 指定给在搜索中访问的第 n 个顶点。证明：这个编号对应着生成树的前序遍历所建立的顶点编号。
45. 用伪码来描述宽度优先搜索的一个变种，它把整数 m 指定给在搜索中访问的第 m 个顶点。
*46. 假设 G 是有向图并且 T 是用宽度优先搜索构造的生成树。证明：G 的每条边都连接同一层的两个顶点、连接一个顶点到低一层的一个顶点或者连接一个顶点到更高层的一个顶点。
47. 证明：如果 G 是有向图并且 T 是用深度优先搜索构造的生成树，则不在这个生成树上的每条边都是连接祖先到后代的*前进边*、连接后代到祖先的*后退边*，或者连接一个顶点到从前访问过的子树的一个顶点的*交叉边*。
*48. 描述深度优先搜索的一个变种，当算法完全处理完一个顶点时，它把最小可用的正整数指定给这个顶点。证明：在这个编号中，每个顶点的编号大于其子女的编号并且子女的编号从左到右递增。

设 T_1 和 T_2 都是一个图的生成树。在 T_1 和 T_2 之间的*距离*是在 T_1 和 T_2 里非 T_1 和 T_2 所共有的边的数目。

49. 求图 9-40 所示图 G 的在图 9-41c 和图 9-42 里所示的每对生成树之间的距离。
*50. 假定 T_1，T_2 和 T_3 都是简单图 G 的生成树。证明：在 T_1 和 T_3 之间的距离不超过 T_1 和 T_2 之间的距离与 T_2 和 T_3 之间的距离的和。
**51. 假定 T_1 和 T_2 都是简单图 G 的生成树。另外，假定 e_1 是在 T_1 里但不在 T_2 里的一条边。证明：存在着在 T_2 里但不在 T_1 里的一条边 e_2，使得若从 T_1 里删除 e_1 而添加 e_2 到 T_1 里，则 T_1 仍然是生成树，并且若从 T_2 里删除 e_2 而添加 e_1 到 T_2 里，则 T_2 仍然是生成树。
*52. 证明：通过相继地删除一条边而添加另外一条边，就有可能从任何一个生成树得出一个生成树的序列。

有向图的*根生成树*是由这个图的边组成的根树，使得这个图的每个顶点都是树里一条边的终点。

53. 对 8.5 节练习 18～23 里的每个有向图来说，求这个图的根生成树，或者确定不存在这样的树。
*54. 证明：每个顶点的入度和出度都相等的连通有向图有根生成树(提示：使用欧拉回路)。
*55. 给出构造每个顶点的入度和出度都相等的连通有向图的根生成树的算法。
*56. 证明：如果 G 是有向图并且 T 是用深度优先搜索构造的生成树，则 G 含有回路当且仅当 G 含有相对于生成树 T 的后退边(参见练习 47)。
*57. 用练习 56 来构造一个确定有向图是否含有回路的算法。

9.5 最小生成树

9.5.1 引言

一个公司计划建立一个通信网络来连接它的五个计算机中心。可以用租用的电话线连接这些中心的任何一对。应当建立哪些连接，以便保证在任何两个计算机中心之间都有通路，且网络的总成本最小？可以用图 9-53 所示的带权图为这个问题建模，其中顶点表示计算机中心，边表示可能租用的电话线，边上的权是边所表示的电话线的月租费。通过找出一棵生成树，使得这棵树的各边的权之和为最小，就可以解决这个问题。这样的生成树称为*最小生成树*。

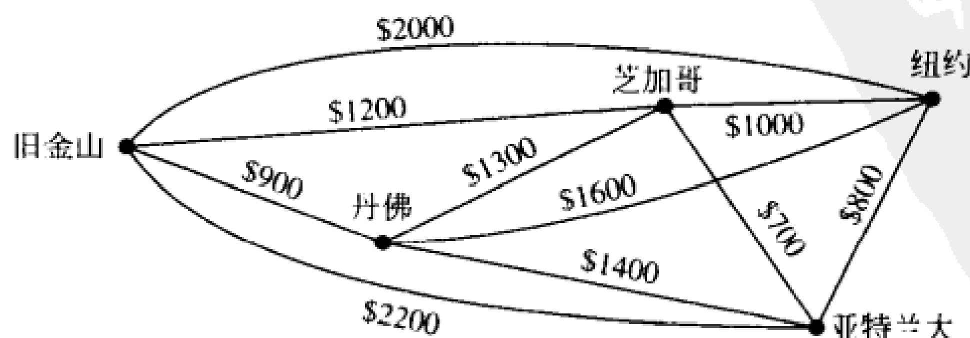


图 9-53 说明计算机网络里的线路的月租费的带权图

9.5.2 最小生成树算法

有大量的问题可以这样解决：求带权图里的一棵生成树，使得这棵树的各边的权之和为最小。

定义1 连通带权图里的最小生成树是具有最小可能的边的权之和的生成树。

下面将给出构造最小生成树的两个算法。这两个算法都是通过添加还没有使用过的、具有规定性质的、权最小的边来进行的。这些算法都是贪心算法的例子。贪心算法是在每个步骤上都做最优选择的算法。在算法的每个步骤上都最优化，并不保证产生全局最优解。不过，本节里给出的构造最小生成树的这两个算法都是产生最优解的贪心算法。

要讨论的第一个算法是罗伯特·普林[⊖]在1957年给出的，虽然这个算法的基本想法有更早的起源。为了执行普林算法，首先选择带最小权的边，把它放进生成树里。相继地向树里添加与已在树里的顶点关联的、并且不与已在树里的边形成简单回路的权最小的边。当已经添加了 $n-1$ 条边时就停止。

本节稍后将证明这个算法产生任何连通带权图的最小生成树。算法1给出普林算法的伪码描述。

```
算法1 普林算法
procedure Prim(G: 带 n 个顶点的连通无向图)
T := 权最小的边
for i := 1 to n-2
begin
  e := 与 T 里顶点关联、并且若添加到 T 里则不形成简单回路的权最小的边
  T := 添加 e 之后的 T
end{T 是 G 的最小生成树}
```

注意，当有超过一条满足相应条件的带相同权的边时，在算法的这个阶段里对所添加的边的选择就不是确定的。需要排序这些边以便让选择是确定的。在本节剩下的部分里将不再考虑这个问题。另外注意，所给的连通带权简单图可能有多于一个的最小生成树(见练习9)。例1和例2说明如何使用普林算法。

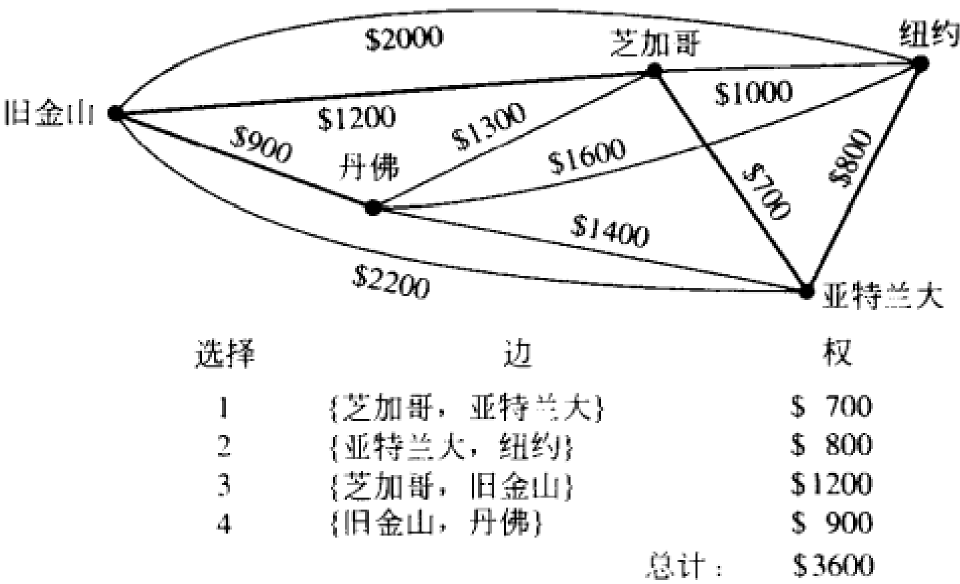
例1 用普林算法设计连接图9-53所表示的所有计算机的、具有最小成本的通信网络。

解 办法是求图9-53的最小生成树。普林算法是这样执行的：选择权最小的初始边，并且相继地添加与树里顶点关联的不形成回路的权最小的边。在图9-54里，加粗的边表示普林算法所产生的最小生成树，并且显示出在每个步骤上所做的选择。■

例2 用普林算法求图9-55所示的图的最小生成树。

解 用普林算法所构造的最小生成树显示在图9-56里。相继选择的边都有显示。■

⊖ **罗伯特·克雷·普林(Robert Clay Prim)**1921年出生在得克萨斯，1941年获得电机工程学士学位，1949年从普林斯顿大学获得数学博士学位。他从1941年到1944年担任通用电气公司的工程师，从1944年到1949年担任美国海军军械实验室的工程师和数学家，从1948年到1949年担任普林斯顿大学的合作研究员。他担任过的其他职务是：1958年到1961年贝尔电话实验室的数学与力学研究主任，以及桑地亚公司的研究副总裁。他目前已经退休。



选择	边	权
1	{芝加哥，亚特兰大}	$ 700
2	{亚特兰大，纽约}	$ 800
3	{芝加哥，旧金山}	$1200
4	{旧金山，丹佛}	$ 900
	总计：	$3600

图 9-54 图 9-53 里带权图的最小生成树

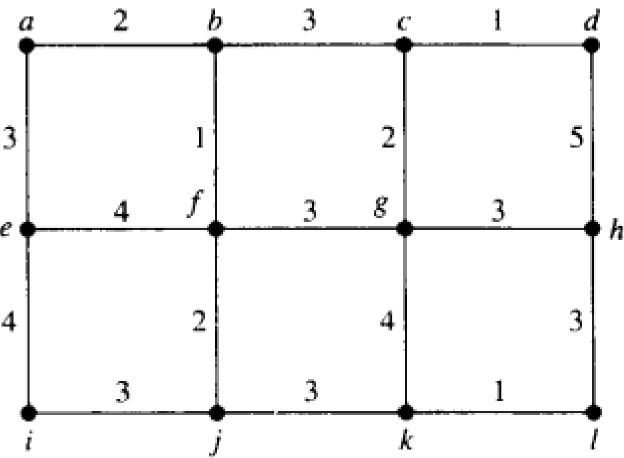
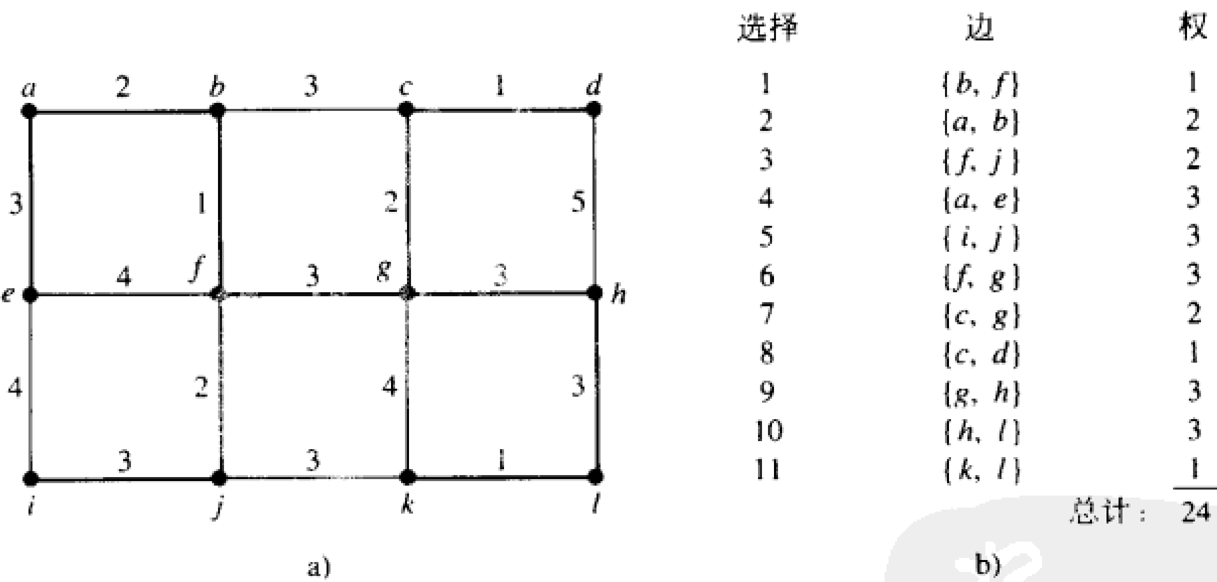


图 9-55 一个带权图



选择	边	权
1	$\{b, f\}$	1
2	$\{a, b\}$	2
3	$\{f, j\}$	2
4	$\{a, e\}$	3
5	$\{i, j\}$	3
6	$\{f, g\}$	3
7	$\{c, g\}$	2
8	$\{c, d\}$	1
9	$\{g, h\}$	3
10	$\{h, l\}$	3
11	$\{k, l\}$	1
	总计：	24

图 9-56 用普林算法构造的最小生成树

将要讨论的第二个算法是约瑟夫·克鲁斯卡尔[⊖]在 1956 年发现的，但在此之前已经有人阐述过这一算法的基本思路。为了执行克鲁斯卡尔算法，要选择图中权最小的一条边。

⊖ **约瑟夫·伯纳德·克鲁斯卡尔(Joseph Bernard Kruskal)**1928 年出生在纽约，后来进入芝加哥大学，1954 年从普林斯顿大学获得博士学位。他是普林斯顿和威斯康星大学的数学教师，随后是密歇根大学的助理教授。1959 年他成为贝尔实验室的技术委员会成员并一直担任这个职务。他目前的研究兴趣包括统计语言学和心理测量学。除了最小生成树的成果之外，克鲁斯卡尔还因为对多维分级的贡献而著名。当克鲁斯卡尔是二年级的研究生时，他发现了产生最小生成树的算法。当时他还不能肯定关于这个题目的两页半的论文是否值得发表，经其他人说服后才递交上去。

相继添加不与已经选择的边形成简单回路的权最小的边。在已经挑选了 $n-1$ 条边之后就停止。

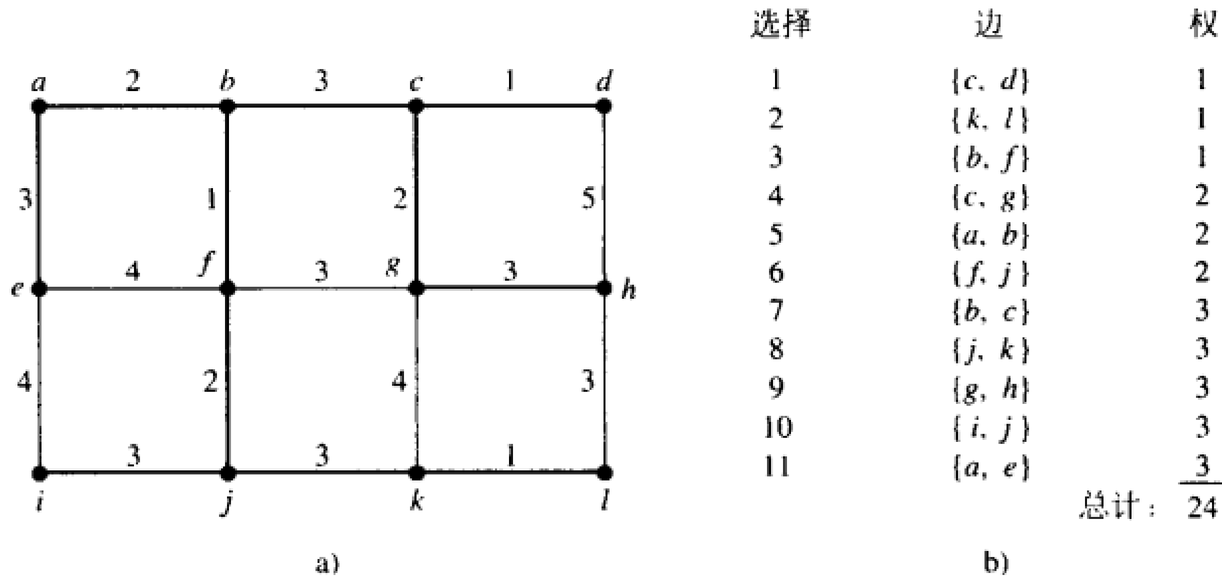
在本节末尾，把证明克鲁斯卡尔算法对每个连通带权图都产生最小生成树留作练习。在算法 2 里给出克鲁斯卡尔算法的伪代码。

```
算法 2  克鲁斯卡尔算法
procedure Kruskal(G: 带 n 个顶点的带权连通无向图)
T := 空图
for i := 1 to n-1
begin
  e := 当添加到 T 里时不形成简单回路的权最小的 G 里的边
  T := 添加 e 之后的 T
end{T 是 G 的最小生成树}
```

读者应当注意普林算法与克鲁斯卡尔算法的区别[⊖]。在普林算法里，选择与已在树里的一个顶点相关联、并且不形成回路的权最小的边；相反地，在克鲁斯卡尔算法里，选择不一定与已在树里的一个顶点相关联、并且不形成回路的权最小的边。注意，在普林算法里，若没有对边排序，则在这个过程的某个阶段上，对添加的边来说就可能有多于一种的选择。因此，为了让这个过程是确定的，就需要对边进行排序。下面的例子说明如何使用克鲁斯卡尔算法。

例 3 用克鲁斯卡尔算法求图 9-55 所示的带权图的最小生成树。

解 在图 9-57 里显示这个最小生成树和在克鲁斯卡尔算法每个阶段上对边的选择。 ■



选择	边	权
1	$\{c, d\}$	1
2	$\{k, l\}$	1
3	$\{b, f\}$	1
4	$\{c, g\}$	2
5	$\{a, b\}$	2
6	$\{f, j\}$	2
7	$\{b, c\}$	3
8	$\{j, k\}$	3
9	$\{g, h\}$	3
10	$\{i, j\}$	3
11	$\{a, e\}$	3
	总计：	24

图 9-57 克鲁斯卡尔算法产生的最小生成树

现在将证明普林算法产生连通带权图的最小生成树。

证 设 G 是一个连通带权图。假定普林算法相继地选择的边是 e_1，e_2，…，e_{n-1}。设 S 是以 e_1，e_2，…，e_{n-1} 作为边的树，而设 S_k 是以 e_1，e_2，…，e_k 作为边的树。设 T 是包含边 e_1，e_2，…，e_k 的 G 的最小生成树，其中 k 是满足下列性质的最大整数：存在着包含普林算法所选择的前 k 条边的最小生成树。若证明了 $S=T$，则该定理得证。

假定 $S\neq T$，所以 $k<n-1$。因此，T 包含边 e_1，e_2，…，e_k，但是不包含 e_{k+1}。考虑由 T

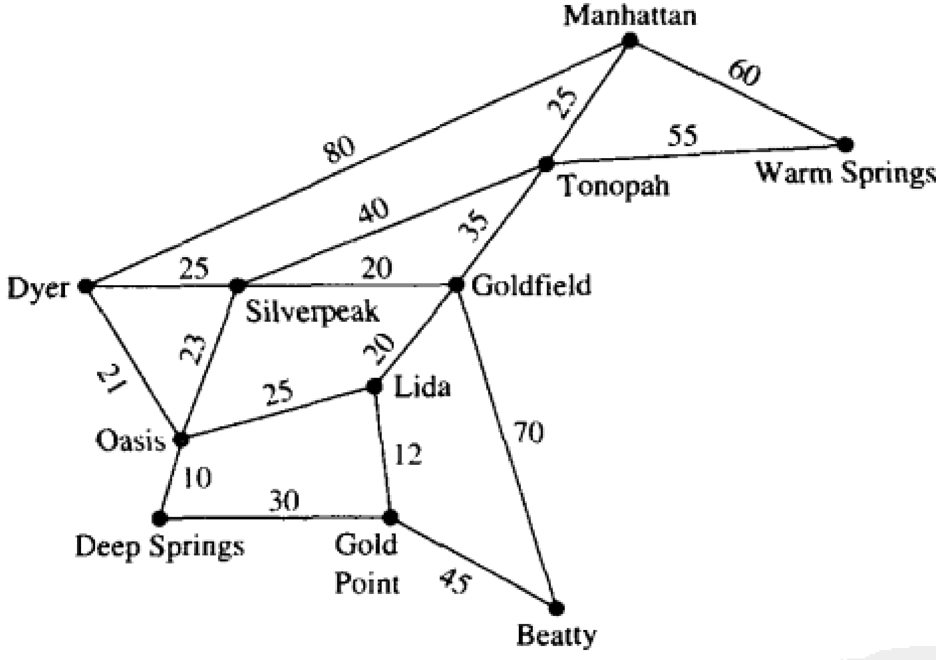
⊖ 历史注解：约瑟夫·克鲁斯卡尔和罗伯特·普林在 20 世纪 50 年代中期提出了构建最小生成树的算法。不过，他们不是第一批发现这些算法的人。例如，人类学家扬·切卡诺夫斯基(Jan Czekanowski)在 1909 年的工作就涵盖了求最小生成树所需要的许多想法。1926 年，奥塔卡·勃鲁乌卡(Otakar Boruvka)在与构造电力网有关的工作中描述了构造最小生成树的方法。

和 e_{k+1} 所组成的图。因为这个图是连通的并且有 n 条边，边过多不可能是树，所以它必然包含简单回路。这个简单回路必然包含 e_{k+1}，因为在 T 里没有简单回路。另外，在这个简单回路里必然有不属于 S_{k+1} 的边，因为 S_{k+1} 是一棵树。通过从 e_{k+1} 的一个端点开始，该端点也是边 e_1，e_2，…，e_k 之一的端点，并且遵循回路直到它到达一条不在 e_{k+1} 里的边为止，就可以找出一条不在 S_{k+1} 里的边 e，它有一个端点也是边 e_1，e_2，…，e_k 之一的端点。通过从 T 里删除 e 并且添加 e_{k+1}，就获得带 $n-1$ 条边的树 T'（它是树，因为它没有简单回路）。注意树 T' 包含 e_1，e_2，…，e_{k+1}。另外，因为普林算法在第 $k+1$ 个步骤上选择 e_{k+1}，并且在这个步骤上 e 也是可用的，所以 e_{k+1} 的权就小于或等于 e 的权。根据这个观察结果就得出 T' 也是最小生成树，因为它的边的权之和不超过 T 的边的权之和。这与对 k 的选择相矛盾，k 是使得包含 e_1，e_2，…，e_k 的最小生成树存在的最大整数。因此，$k=n-1$ 并且 $S=T$。所以普林算法产生最小生成树。 □

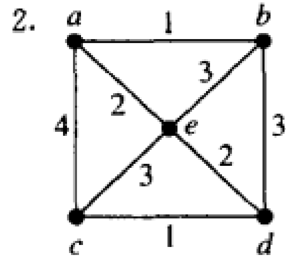
可以证明（参见[CoLeRiSt01]）为了求出具有 e 条边和 v 个顶点的图的最小生成树，可以用 $O(e \log e)$ 次运算来完成克鲁斯卡尔算法并且可以用 $O(e \log v)$ 次运算来完成普林算法。因此，对于稀疏图来说使用克鲁斯卡尔算法更好。在稀疏图中，e 远远小于 $C(v, 2)=v(v-1)/2$，即具有 n 个顶点的无向图的可能的总边数。否则，这两个算法的复杂性没有什么差别。

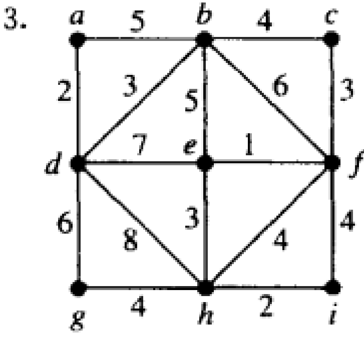
练习

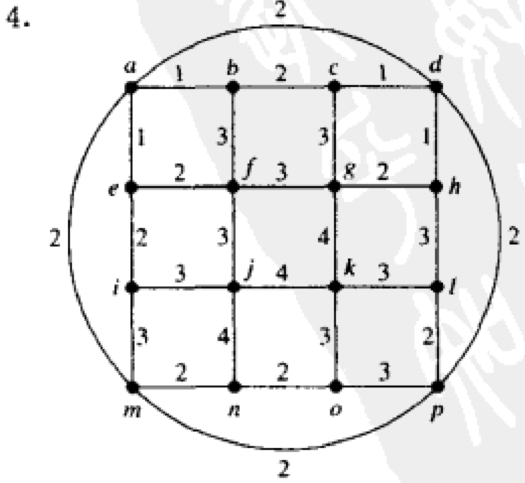
1. 下图所表示的道路都还没有铺设路面。边的权表示在成对的乡镇之间的道路长度。哪些道路应当铺设路面，以便在每对乡镇之间都有铺设路面的道路，而且使得铺设的道路的长度最短？（注：这些乡镇都在内华达州。）



在练习 2～4 里，用普林算法求所给的带权图的最小生成树。

2. 

3. 

4. 

5. 用克鲁斯卡尔算法设计在本节开头所描述的通信网络。
6. 用克鲁斯卡尔算法求练习2里带权图的最小生成树。
7. 用克鲁斯卡尔算法求练习3里带权图的最小生成树。
8. 用克鲁斯卡尔算法求练习4里带权图的最小生成树。
9. 找出具有多于一棵最小生成树的、带有最少可能边数的连通带权简单图。
10. 带权图里的*最小生成森林*是权最小的生成森林。解释如何修改普林算法和克鲁斯卡尔算法来构造最小生成森林。

连通带权无向图的*最大生成树*是带最大可能的权的生成树。

11. 设计与普林算法类似的、构造连通带权图的最大生成树的算法。
12. 设计与克鲁斯卡尔算法类似的、构造连通带权图的最大生成树的算法。
13. 求练习2里带权图的最大生成树。
14. 求练习3里带权图的最大生成树。
15. 求练习4里带权图的最大生成树。
16. 求在本节开头所提出问题里连接五个计算机中心的次最便宜的通信网络。

*17. 设计求连通带权图里次最短生成树的算法。
*18. 证明：连通带权图里权最小的边，必然属于任何一个最小生成树。
19. 证明：若所有边的权都不相同，则连通带权图里有唯一的最小生成树。
20. 假定连接图9-53里城市的计算机网络必须包含纽约与丹佛之间的直接连接。那么应当包含哪些其他的连接，使得在每两个计算机中心之间都存在连接，并且费用最少？
21. 求图9-55的带权图里包含边$\{e, i\}$和$\{g, k\}$的总权最小的生成树。
22. 描述一个算法，它求连通带权无向简单图里包含所规定的一组边的权最小的生成树。
23. 用伪码表达练习22里设计的算法。

索林(Sollin)*算法*从连通带权简单图$G=(V, E)$这样产生最小生成树：相继地添加成组的边。假定对V里顶点进行了排序。这样就产生边的一个顺序，其中若u_0先于u_1，或者若$u_0=u_1$并且v_0先于v_1，则$\{u_0, v_0\}$先于$\{u_1, v_1\}$。这个算法首先同时选择每个顶点关联的权最小的边。在平局情形下选择在上述顺序里的第一条边。这样就产生出一个没有简单回路的图，即一些树组成的一个森林(练习24要求证明这个事实)。其次，对森林里的每棵树，同时选择在该树里一个顶点与在不同的一棵树里顶点之间的最短的边。同样在平局情形下选择在上述顺序里的第一条边。(这样就产生出一个没有简单回路的图，它包含着比在这一步之前出现的更少的树；参见练习24。)继续进行同时地添加连接树的边的过程，直到已经选择了$n-1$条边为止。在这个阶段上就已经构造了一棵最小生成树。

*24. 证明：在索林算法的每个阶段上边的添加都产生一个森林。
25. 用索林算法产生图9-53和图9-55所示的带权图的最小生成树。
*26. 用伪代码表达索林算法。
**27. 证明：索林算法产生连通无向带权图里的最小生成树。
*28. 证明：索林算法的第一步产生至少包含$\lceil n/2 \rceil$条边的森林。
*29. 证明：若在索林算法的某个中间步骤里存在r棵树，则算法的下一次迭代至少添加$\lceil n/2 \rceil$条边。
*30. 证明：在已经完成索林算法的第一步、并且已经$k-1$次执行索林算法的第二步之后，还剩下不超过$\lceil n/2^k \rceil$棵树。
*31. 证明：索林算法至少需要$\log n$次迭代，以便从带有n个顶点的连通无向带权图产生一棵最小生成树。
32. 证明：克鲁斯卡尔算法产生最小生成树。

关键术语和结果

术语

树：没有简单回路的连通无向图。

森林：没有简单回路的无向图。

根树：具有一个规定的顶点(称为根)，使得从这个根到任意其他顶点有唯一通路的有向图。

子树：树的子图，子图本身也是一棵树。

根树里 v 的父母：使得(u, v)是根树的一条边的顶点 u。

根树里顶点 v 的子女：以 v 作为父母的任何顶点。

根树里顶点 v 的兄弟：与 v 具有相同父母的顶点。

根树里顶点 v 的祖先：在从根到 v 的通路上的任何顶点。

根树里顶点 v 的后代：以 v 作为祖先的任何顶点。

内点：具有子女的顶点。

树叶：没有子女的顶点。

顶点的层：从根到这个顶点的通路的长度。

树的高度：树里顶点的最大层数。

m 元树：每个内点都有不超过 m 个子女的树。

满 m 元树：每个内点都有恰好 m 个子女的树。

二叉树：满足 $m=2$ 的 m 元树(可以指定每个子女作为父母的左子女或右子女)。

有序树：在其中对每个内点的子女都线性地排序的树。

平衡树：在其中每个顶点都是在 h 层或 $h-1$ 层上的树，其中 h 是这棵树的高度。

二叉搜索树：这样的二叉树，在其中以项对顶点进行标记，使得一个顶点的标记大于这个顶点的左子树里所有顶点的标记，并且小于这个顶点的右子树里所有顶点的标记。

决策树：性质如下的根树，在其中每个顶点表示一次决策的可能输出，而树叶表示可能的解。

前缀码：一种编码，其中一个字符的编码永远不是另外一个字符的编码的前缀。

最小最大策略：第一个选手和第二个选手分别移动到具有最大值和最小值的顶点所表示的局面的策略。

博弈树里顶点的值：对于树叶来说，就是当游戏在这个树叶所表示的局面里结束时第一个选手的得分。对于分别在偶数或奇数层上的内点来说，就是这个内点的子女的最大值或最小值。

树的遍历：树的顶点的列表。

前序遍历：通过规定列出根，接着列出第一棵子树，接着以从左到右的出现顺序列出其余子树，来递归地定义的有序根树的顶点列表。

中序遍历：通过规定列出第一棵子树，接着列出根，接着以从左到右的出现顺序列出其余子树，来递归地定义的有序根树的顶点列表。

后序遍历：通过规定以从左到右的出现顺序列出各子树，接着列出根，来递归地定义的有序根树的顶点列表。

中缀记法：从表示表达式(包括全套括号)的二叉树的中序遍历所获得的表达式形式。

前缀记法，或波兰记法：从表示表达式的二叉树的前序遍历所获得的表达式形式。

后缀记法，或逆波兰记法：从表示表达式的二叉树的后序遍历所获得的表达式形式。

生成树：包含图的所有顶点的树。

最小生成树：带最小可能的边的权之和的生成树。

贪心算法：在每个步骤上都做出最优选择的最优化算法。

结果

一个图是树，当且仅当在它的任何两个顶点之间都存在唯一简单通路。

带有 n 个顶点的树具有 $n-1$ 条边。

带有 i 个内点的满 m 元树具有 $mi+1$ 个顶点。

在满 m 元树的顶点数、树叶数和内点数之间的关系(见9.1节定理4)。

在高度为 h 的满 m 元树里至多有 m^h 个树叶。

若 m 元树有 l 个树叶，则它的高度至少是 $\lceil \log_m l \rceil$。若这树也是满的和平衡的，则它的高度就是 $\lceil \log_m l \rceil$。

赫夫曼编码：给定一组符号的频率，为这些符号构造最优二元码的过程。

深度优先搜索，或回溯：构造生成树的过程，通过添加形成通路的边，直到不可能这样做为止，然后沿这条通路往回移动，直到找到可以形成新的通路的顶点为止。

宽度优先搜索：构造生成树的过程，通过相继添加与上次添加的边相关联的所有边，除非形成简单回路。

普林算法：产生带权图里最小生成树的过程，通过相继添加与已经在树里的顶点相关联的所有边中权最小的边，使得再添加边时不会产生简单回路。

克鲁斯卡尔算法：产生带权图里最小生成树的过程，通过相继添加还不在树里的权最小的边，使得再添加边时不会产生简单回路。

复习题

1. a)定义树。　　　b)定义森林。
2. 在树的顶点之间能否有两条不同的简单通路?
3. 至少给出三个例子说明如何在建模中使用树。
4. a)定义根树和这样的树的根。
   b)定义根树里顶点的父母和顶点的子女。
   c)什么是根树里的内点、树叶和子树?
   d)画出至少带10个顶点的根树，其中每个顶点的度都不超过3。指出树根、每个顶点的父母、每个顶点的子女、内点和树叶。
5. a)带 n 个顶点的树有多少条边?
   b)为确定带有 n 个顶点的森林里的边数，你需要知道什么值?
6. a)定义正则 m 元树。
   b)若正则 m 元树有 i 个内点，则它有多少个顶点?此树有多少个树叶?
7. a)什么是根树的高度?
   b)什么是平衡树?
   c)高度为 h 的 m 元树可以有多少个树叶?
8. a)什么是二叉搜索树?
   b)描述构造二叉搜索树的算法。
   c)构造单词 *vireo*，*warbler*，*egret*，*grosbeak*，*nuthatch* 和 *kingfisher* 的二叉搜索树。
9. a)什么是前缀码?
   b)二叉树如何表示前缀码?
10. a)定义前序遍历、中序遍历和后序遍历。
    b)给出至少带12个顶点的二叉树的前序遍历、中序遍历和后序遍历的实例。

11. a)解释一下如何用前序遍历、中序遍历和后序遍历来求算术表达式的前缀形式、中缀形式和后缀形式。
    b)画出表示$((x-3)+((x/4)+(x-y)\uparrow 3))$的有序根树。
    c)求在(b)里的表达式的前缀和后缀形式。
12. 证明：排序算法所使用的比较次数至少是$\lceil \log n! \rceil$。
13. a)描述一下赫夫曼编码算法，这个算法求一组给定频率的符号的最优编码。
    b)用赫夫曼编码求下列符号和频率的最优编码：A：0.2，B：0.1，C：0.3，C：0.4。
14. 画出 nim 游戏的博弈树，假设初始局面由两堆石头组成，分别含有 1 块和 4 块石头。假如两个选手都遵循最优策略的话，谁将赢得游戏呢?
15. a)什么是简单图的生成树?
    b)哪些简单图具有生成树?
    c)描述需要求出简单图的生成树的应用，至少举两个不同的应用。
16. a)描述求简单图里生成树的两个不同算法。
    b)用你所选择的至少带 8 个顶点和 15 条边的图，来解释你在(a)里所描述的两个算法是如何求简单图的生成树的。
17. a)解释如何用回溯来确定能否用 n 种颜色来着色简单图。
    b)用例子说明如何用回溯来证明：色数等于 4 的图不能用 3 种颜色来着色，但是可以用 4 种颜色来着色。
18. a)什么是连通带权图的最小生成树?
    b)至少描述出两个不同的、需要求出连通带权图的最小生成树的应用。
19. a)描述求最小生成树的普林算法和克鲁斯卡尔算法。
    b)用至少带 8 个顶点和 15 条边的图，来解释克鲁斯卡尔算法和普林算法是如何求最小生成树的。

补充练习

*1. 证明：简单图是树当且仅当它不包含简单回路，并且添加连接两个不相邻顶点的一条边，就产生恰好有两条回路的新图(其中不认为包含相同的边的回路是不同的)。
*2. 有多少种非同构的带 6 个顶点的根树?
3. 证明：每一个至少有一条边的树都至少有两个悬挂点。
4. 证明：有 $n-1$ 个悬挂点的、带有 n 个顶点的树必然同构于 $K_{1,n-1}$。
5. 带有 n 个顶点的树的顶点的度之和是什么?
*6. 假定 d_1，d_2，…，d_n 是和为 $2n-2$ 的 n 个正整数。证明：存在一个带有 n 个顶点的树，使得这些顶点的度为 d_1，d_2，…，d_n。
7. 证明：每个树都是可平面图。
8. 证明：每个树都是偶图。
9. 证明：每个森林都可以用两种颜色来着色。

k 度 B 树是一个根树，它的所有树叶都是在同一层上，它的根具有至少两个和至多 k 个子女，除非根就是树叶，并且除根外的每个内点有至少$\lceil k/2 \rceil$个、但不超过 k 个子女。当用 B 树来表示计算机文件时，就可以有效地访问这些文件。

10. 画出三种不同的高度为 4 的 3 度 B 树。
*11. 给出高度为 h 的 k 度 B 树里树叶数的上界和下界。
*12. 给出有 n 个树叶的 k 度 B 树的高度的上界和下界。

二项式树 $B_i(i=0, 1, 2, \cdots)$是如下递归定义的有序根树。

基础步骤：二项式树 B_0 是具有单个顶点的树。

递归步骤：设 k 是非负整数。为了构造二项式树 B_{k+1}，把 B_k 的一个副本加入 B_k 的第二个副本，方法

是加入一条边，这条边让 B_k 的第一个副本的根成为 B_k 的第二个副本的最左子女。

13. 对 $k=0$，1，2，3，4，画出 B_k。
14. B_k 有多少个顶点？证明你的答案是正确的。
15. 求 B_k 的高度。证明你的答案是正确的。
16. B_k 在深度 j 有多少个顶点？其中 $0 \leqslant j < k$。证明你的答案是正确的。
17. B_k 的根的度数是多少？证明你的答案是正确的。
18. 证明：B_k 中度数最大的顶点是根。

若根树 T 满足下面的递归定义，则称它为 S_k 树。若它只有一个顶点，则它是 S_0 树。对 $k>0$ 来说，若通过把一个 S_{k-1} 树的根作为一个新树的根，把另外一个 S_{k-1} 树的根作为新树的根的子女，从两个 S_{k-1} 树来建立一个新树，则这个新树是 S_k 树。

19. 对 $k=0$，1，2，3，4，画出一个 S_k 树。
20. 证明：S_k 树有 2^k 个顶点并且在 k 层上有唯一一个顶点。在 k 层上的这个顶点称为**把柄**。

*21. 假定 T 是带有把柄 v 的 S_k 树。证明：T 可以从不相交的树 T_0，T_1，…，T_{k-1} 来这样获得，其中 v 不在这些树的任何一个里，对 $i=0$，1，…，$k-1$ 来说，T_i 是 S_i 树；对 $i=0$，1，…，$k-2$，连接 v 到 v_0 并且连接 r_i 到 r_{i+1}。

有序根树在**层顺序**下的顶点列表从根开始，接着是从左到右在 1 层上的顶点，从左到右在 2 层上的顶点，依此类推。

22. 列出 9.3 节图 9-27 和图 9-33 里的有序根树在层顺序下的顶点列表。
23. 设计列出有序根树在层顺序下的顶点列表的算法。

*24. 设计一种算法，确定一组通用地址能否成为根树的树叶地址。

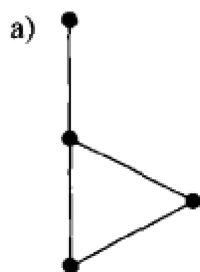
25. 设计从树叶的通用地址来构造根树的算法。

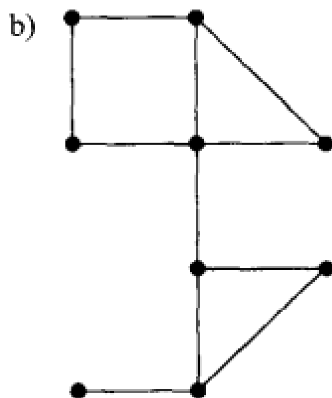
图的**割集**是这样一些边的集合，使得删除这些边就产生一个子图，这个子图的连通分支比原来的图要多，但是这些边的任何真子集都没有这个性质。

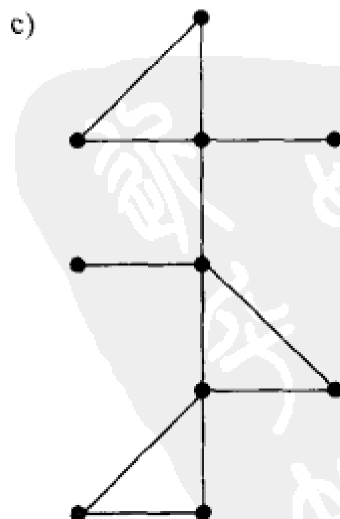
26. 证明：图的割集必然与这个图的任何生成树都有至少一条公共边。

仙人掌图是连通图，其中没有边是在多于一条的简单回路上，这些简单回路不经过除了起点以外的任何顶点超过一次，而且不在除了终点以外的其他地方经过起点(其中不认为由相同的边所组成的两个回路是不同的)。

27. 下面的图哪些是仙人掌图？

a)


b)


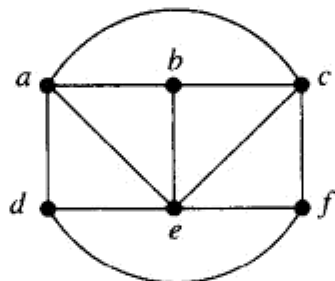
c)


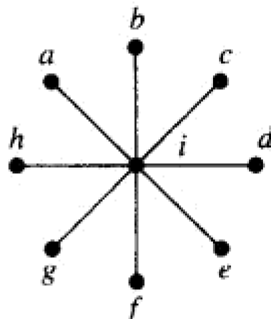
28. 树是否必然是仙人掌图？
29. 证明：若在树里添加一条回路，它包含一些起止于树里顶点上的新边，则形成一个仙人掌图。

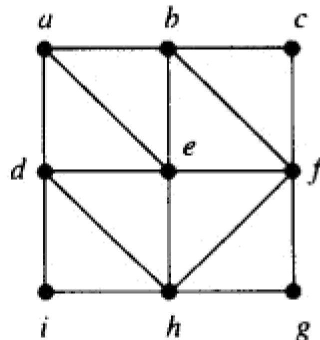
*30. 证明：若在连通图里，每条不经过任何顶点(除起点以外)超过一次的回路都包含奇数条边，则这个图必然是仙人掌图。

简单图 G 的限制度数生成树具有下面性质，在这个树里顶点的度不能超过某个规定的界限。在运输系统的模型里，在交叉路口处的道路数目是有限的，在通信网络的模型里，进入一个结点的连接数目是有限的，这种模型都可以使用限制度数生成树。

在练习 31～33 里，求所给的图的限制度数生成树，其中每个顶点的度都小于或等于 3，或者证明不存在这样的生成树。

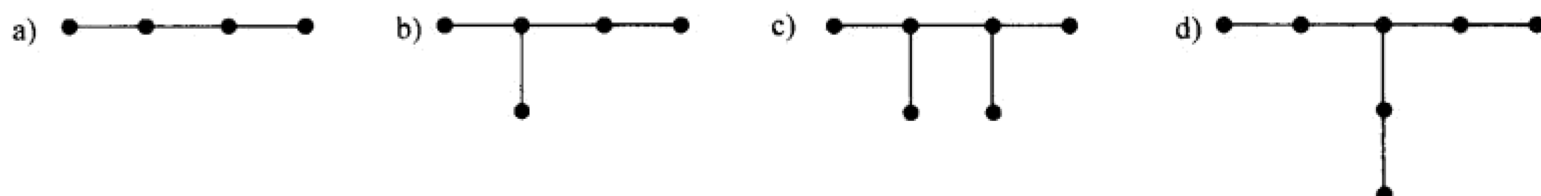
31.


32.


33.


34. 证明：简单图的限制度数生成树(其中每个顶点的度都不超过 2)是由该图里的单独一条哈密顿回路所组成的。

35. 若可以用整数 1，2，…，n 来标记带有 n 个顶点的树的顶点，使得相邻顶点的标记之差的绝对值全都是不同的，则称这棵树为*优美的*。证明：下面的树都是优美的。



*毛虫图*是含有一条简单通路的树，使得不包含在这条通路里的每个顶点都与这条通路里的一个顶点相邻。

36. 练习 35 里的图，哪些是毛虫图？

37. 带 6 个顶点的互不同构的毛虫图有多少种？

**38. a)证明或反驳：其边形成一条简单通路的所有树都是优美的。

b)证明或反驳：所有的毛虫图都是优美的。

39. 假设在一个很长的位串中，位 0 出现的频率是 0.9，位 1 出现的频率是 0.1，并且位都是独立地出现的。

a)为 4 个两位的块 00、01、10 和 11 构造赫夫曼编码。用这个编码来编码一个位串，所需要的平均位数是多少？

b)为 8 个三位的块构造赫夫曼编码。用这个编码来编码一个位串，所需要的平均位数是多少？

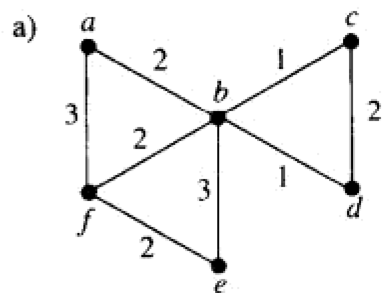
40. 假设 G 是没有回路的有向图。描述如何用深度优先搜索来完成 G 的顶点的拓扑排序。

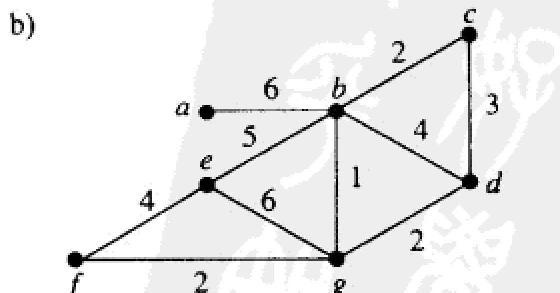
*41. 假定 e 是带权图里与顶点 v 关联的一条边，使得 e 的权不超过与顶点 v 关联的任何其他边的权。证明：存在一棵包含这条边的最小生成树。

42. 三对夫妇到达一条河流的岸边。每个妻子都容易嫉妒，当她的丈夫与其他的妻子(或其他人)在一起、但是她不在场时，她就不信任她的丈夫。六个人如何用一条只能装载不超过两个人的船来渡河，使得每位丈夫无法与妻子之外的女人单独相处？解答时使用图论模型。

*43. 证明：若在带权图里没有两条边具有相同的权，则在每个最小生成树里都包含着与顶点 v 关联的权最小的边。

44. 求下面两个图的最小生成树，其中在生成树里每个顶点的度都不超过 2。

a)


b)


计算机题目

按下述给定的输入与输出写出程序。

1. 给定无向简单图的相邻矩阵，确定这个图是不是树。
2. 给定根树的相邻矩阵和这棵树里的一个顶点，求出这个顶点的父母、子女、祖先、后代和层数。
3. 给定根树的边的列表和这棵树里的一个顶点，求出这个顶点的父母、子女、祖先、后代和层数。
4. 给定项的列表，构造包含这些项的二叉搜索树。
5. 给定二叉搜索树和一个项，在这个二叉搜索树里求出这个项的位置或添加这个项。
6. 给定有序根树的边的有序列表，求出它的边的通用地址。
7. 给定有序根树的边的有序列表，以前序、中序和后序列出它的顶点。
8. 给定前缀形式的算术表达式，求它的值。
9. 给定后缀形式的算术表达式，求它的值。
10. 给定一组符号的频率，用赫夫曼编码来求这些符号的最优编码。
11. 给定轮流取石头游戏的开局，确定第一个选手的最优策略。
12. 给定连通无向简单图的相邻矩阵，用深度优先搜索找出这个图的生成树。
13. 给定连通无向简单图的相邻矩阵，用宽度优先搜索找出这个图的生成树。
14. 给定一组正整数和一个正整数 N，利用回溯求这些整数的其和为 N 的子集合。

*15. 给定无向简单图的相邻矩阵，若有可能，则利用回溯来用三种颜色为这个图着色。

*16. 给定一个正整数 n，利用回溯来解决 n 皇后问题。

17. 给定带权无向连通图的边的列表和它们的权，用普林算法求这个图的最小生成树。
18. 给定带权无向连通图的边的列表和它们的权，用克鲁斯卡尔算法求这个图的最小生成树。

计算和研究

利用计算程序或你所编写的程序来做下面的练习。

1. 显示所有的带有 6 个顶点的树。
2. 显示全部的互不同构的带有 7 个顶点的树。

*3. 根据 ASCⅡ码字符在典型输入中出现的频率，构造它们的赫夫曼编码。

4. 对 $n=1, 2, 3, 4, 5, 6$，计算 K_n 的不同的生成树的个数。猜想它们的公式。
5. 对于 $n=100$，$n=1000$ 和 $n=10\,000$ 来说，比较一下排序 n 个元素的表所需要的比较次数，其中元素都是随机选择的正整数，使用选择排序、插入排序、归并排序和快速排序。
6. 对于不超过 10 的所有正整数 n 来说，计算一下在 $n\times n$ 棋盘上放置 n 个皇后，使得这些皇后不能互相攻击的不同方式数。

*7. 求把美国 50 个州的首府互相连接起来的图的最小生成树，其中每条边上的权是首府之间的距离。

8. 画出 4×4 棋盘上跳棋游戏的完全博弈树。

写作题目

利用本书以外的资料，就下列问题写出短文。

1. 解释凯莱如何用树来枚举特定类型的碳水化合物的个数。
2. 定义 AVL 树(有时也称为均高树)。解释如何以及为什么在许多不同的算法中都用到 AVL 树。
3. 定义四叉树并解释如何用四叉树来表示图像。描述如何通过操纵对应的四叉树来旋转、缩放和转换图像。
4. 定义一个堆，并解释如何把树转化成堆。堆为什么在排序中有用？
5. 描述一下针对连续读入字符时字母频率发生变化的数据压缩的动态规划算法，比如自适应赫夫曼编码。
6. 解释如何用 $\alpha-\beta$ 剪枝来简化对博弈树的值的计算。

7. 描述下棋程序(比如深蓝)所使用的技术。
8. 为树网这种类型的图下定义。解释一下这种图如何用在非常大的系统集成和并行计算中。
9. 讨论在 IP 组播中避免路由器之间回路所用的算法。
10. 描述基于深度优先搜索来求图的节点的算法。
11. 描述基于深度优先搜索来求有向图的强连通分支的算法。
12. 描述在 Web 上不同搜索引擎的网络爬虫和网络蜘蛛所用的搜索技术。
13. 描述一下求图的最小生成树的算法，使得生成树上任意顶点的最大度数不超过一个固定的常数 k。
14. 就复杂度和应用场合而言，对一些最重要的排序算法进行比较。
15. 讨论一下构造最小生成树的算法的历史和起源。
16. 描述产生随机树的算法。



第10章 布尔代数

计算机和其他电子设备中的电路都有输入和输出，输入是0或1，输出也是0或1。电路可以用任何具有两个不同状态的基本元件来构造，开关和光学装置都是这样的元件。开关可能处于开或关的位置，光学装置可能是点亮或未点亮的。1854年，乔治·布尔(George Boole)在《*The Laws of Thought*》一书中第一次给出了逻辑的基本规则。1938年，克劳德·香农(Claude Shannon)⊖揭示了怎么用逻辑的基本规则来设计电路，这些基本规则形成了布尔代数的基础。本章将逐步展开布尔代数基本性质的讨论。电路的操作可以用布尔函数来定义，这样的布尔函数对任意一组输入都能指出其输出的值。构造电路的第一步是用由布尔代数的基本运算构造的表达式来表示布尔函数。我们将介绍一个能产生这些表达式的算法，所得到的表达式可能包含一些冗余运算。本章的后面部分将描述一个求表达式方法，求得的表达式中所包含的和与积的个数是表示一个布尔函数所需数量中最少的。将要展开讨论的这些方法称为卡诺(Karnaugh)图方法和奎因-莫可拉斯基(Quine-McCluskey)方法，它们对于有效电路的设计十分重要。

10.1 布尔函数

10.1.1 引言

布尔代数提供的是集合{0，1}上的运算和规则，这个集合及布尔代数的规则还可以用来研究电子和光学开关。我们将来用得最多的三个布尔代数运算是补、布尔和与布尔积。元素的补用上划线加以标记，其定义为：$\overline{0}=1$ 且 $\overline{1}=0$。布尔和记为+或 *OR*，它的值如下：

$$1+1=1,\quad 1+0=1,\quad 0+1=1,\quad 0+0=0$$

布尔积记为·或 *AND*，它的值如下：

$$1\cdot 1=1,\quad 1\cdot 0=0,\quad 0\cdot 1=0,\quad 0\cdot 0=0$$

在不引起混淆时，可以删去“·”，就像在写代数积时所做的那样。除非使用括号，布尔运算的优先级规则是：首先计算所有补，然后是布尔积，然后是布尔和，如例1所示。

例1 计算 $1\cdot 0+\overline{(0+1)}$ 的值。

解 根据补、布尔积与布尔和的定义得

⊖ **克劳德·艾尔伍德·香农(Claude Elwood Shannon, 1916—2001)** 香农1916年生于密歇根州的盖劳得(Gaylod)，1936年毕业于密歇根大学，之后在麻省理工学院继续学习。在麻省理工学院，他得到了一份维护微分分析器的工作。这种机器是由轴和齿轮构成的机械计算装置，由他的硕士论文导师文那瓦·布什(Vannevar Bush)构造。这篇硕士论文写于1936年，研究微分分析器的逻辑性质，第一次提出了布尔代数在开关电路设计中的应用，可能是20世纪最著名的硕士论文。他于1940年从麻省理工学院获得博士学位，并于1940年加入贝尔实验室。在这里他从事数据有效传输方面的工作。他是首批用数位表示信息的人之一。在贝尔实验室，他还从事确定电话线所能承载的流量方面的工作。香农对信息论做出了许多十分重要的贡献。在20世纪50年代早期，他是人工智能的奠基人之一。他在1956年加入麻省理工学院继续从事信息论研究。

香农有着不从常规的一面。他被认为是以火箭为动力的塑料玩具飞盘的发明者。他曾在贝尔实验室的门厅里骑着单轮脚踏车，并同时耍着4个球，且因此而闻名。香农50岁就退休了，但在其后的十多年中还零星地发表文章。他在晚年从事一些宠物计划，如建造一个用于跳跃的摩托化高跷杖。香农的一句有意思的语录是：“我可以想象，我们将成为机器人，狗将成为人的时刻将会到来，我为机器鼓气加油”。这句话于发表在1987年的《*Omni Magazine*》杂志上。

$$1\cdot 0+\overline{(0+1)}^{①}=0+\overline{1}$$
$$=0+0$$
$$=0$$ ■

补、布尔积和布尔和分别对应于逻辑运算¬、∨和∧，且0对应于F(假)，1对应于T(真)。关于布尔代数的结果可以直接翻译成关于命题的结果。相反地，关于命题的结果也能翻译成关于布尔代数的命题。

10.1.2 布尔表达式和布尔函数

设 $B=\{0,1\}$，则 $B^n=\{(x_1,x_2,\cdots,x_n)\mid x_i\in B,1\leqslant i\leqslant n\}$ 是由0和1所能构成的所有 n 元有序列的集合。变元 x 如果仅从 B 中取值，则称该变元为*布尔变元*，即它的值只可能为0或1。从 B^n 到 B 的函数称为 *n 度布尔函数*[⊖]。

例2 从布尔变元有序对之取值集合到集合 $\{0,1\}$ 的函数 $F(x,y)=x\overline{y}$ 就是一个2度布尔函数，且 $F(1,1)=0$，$F(1,0)=1$，$F(0,1)=0$，$F(0,0)=0$。F 的值如表10-1所示。 ■

表 10-1

x	y	$F(x,y)$
1	1	0
1	0	1
0	1	0
0	0	0

布尔函数也可用由变元和布尔运算构成的表达式来表示。关于变元 $x_1,x_2,\cdots,x_n$ 的*布尔表达式*递归定义如下：

1)0，1，$x_1,x_2,\cdots,x_n$ 是布尔表达式；

2)如果 E_1 和 E_2 是布尔表达式，则 $\overline{E_1}$，(E_1E_2) 和 (E_1+E_2) 是布尔表达式。

每个布尔表达式都表示一个布尔函数，此函数的值是通过在表达式中用0和1替换变元得到的。在10.2节中我们将介绍怎么用布尔表达式来表示布尔函数。

例3 求由 $F(x,y,z)=xy+\overline{z}$ 表示的布尔函数的值。

解 这个函数的值由表10-2表示。 ■

表 10-2

x	y	z	xy	$\overline{z}$	$F(x,y,z)=xy+\overline{z}$
1	1	1	1	0	1
1	1	0	1	1	1
1	0	1	0	0	0
1	0	0	0	1	1
0	1	1	0	0	0
0	1	0	0	1	1
0	0	1	0	0	0
0	0	0	0	1	1

注意，布尔函数还可用图形来表示，方法是：将 n 元二进制数组与 n 方体的顶点一一对应，再标出那些函数值为1的顶点。

例4 例3中从 B^3 到 B 的函数 $F(x,y,z)=xy+\overline{z}$ 可如下表示：标出满足 $F(x,y,z)=1$ 的五个3元组(1，1，1)、(1，1，0)、(1，0，0)、(0，1，0)和(0，0，0)所对应的顶点。如图10-1所示，这些顶点用实心的黑圈标出。 ■

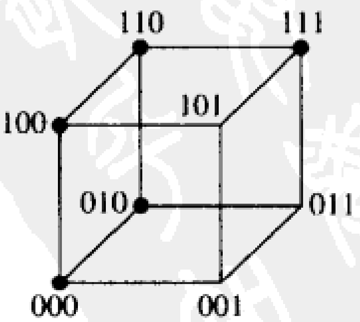


图 10-1

n 个变元的布尔函数 F 和 G 是相等的，当且仅当 $F(b_1,b_2,\cdots,b_n)=G(b_1,b_2,\cdots,b_n)$，其中 $b_1,b_2,\cdots,b_n$ 属于 B。

⊖ 原文为 Boolean function of degree n，常称为 n 元布尔函数。——译者注

表示同一个函数的不同的布尔表达式称为是等价的。例如，布尔表达式 xy、$xy+0$ 和 $xy\cdot 1$ 都是等价的。布尔函数 F 的补函数是 $\overline{F}$，此处 $\overline{F}(x_1,\cdots,x_n)=\overline{F(x_1,\cdots,x_n)}$。设 F 和 G 是 n 度的布尔函数，函数的布尔和 $F+G$ 与布尔积 FG 分别定义为

$$(F+G)(x_1,\cdots,x_n)=F(x_1,\cdots,x_n)+G(x_1,\cdots,x_n)$$
$$(FG)(x_1,\cdots,x_n)=F(x_1,\cdots,x_n)G(x_1,\cdots,x_n)$$

2 度布尔函数是从一个 4 个元素的集合到 B 的函数，这 4 个元素是 $B=\{0,1\}$ 中元素构成的元素对，B 是有 2 个元素的集合。因而有 16 个不同的 2 度布尔函数。在表 10-3 中，我们列出了这 16 个 2 度布尔函数的值，这 16 个不同的 2 度布尔函数被记为 F_1，F_2，…，F_{16}。

表 10-3　2 度布尔函数

x	y	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	F_{14}	F_{15}	F_{16}
1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0	0	0
1	0	1	1	1	1	0	0	0	0	1	1	1	1	0	0	0	0
0	1	1	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0
0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0

例 5　有多少个不同的 n 度布尔函数?

解　由计数的乘积规则知：有 2^n 个由 0 和 1 构成的不同的 n 元组。因为布尔函数就是对这 2^n 个 n 元组中的每一个进行赋值，故乘积规则表明有 2^{2^n} 个不同的 n 度布尔函数。 ■

表 10-4 列出了 1～6 度不同布尔函数的个数。这种函数的个数增长非常快。

表 10-4　n 度布尔函数的个数

度数	数量	度数	数量
1	4	4	65 536
2	16	5	4 294 967 296
3	256	6	18 446 744 073 709 551 616

10.1.3　布尔代数恒等式

布尔代数有许多恒等式，表 10-5 列出了其中最重要的恒等式。(读者应将这些恒等式与 1.2 节的表 1-13 中的逻辑等价以及 1.7 节的表 1-21 中的集合恒等式进行比较，所有这些都是一个更抽象结构中恒等式集合的特殊情形。)这些恒等式对于电路设计的简化特别有用。表 10-5 中的每个恒等式都可以用表来证明。下面的例子就以这种方法证明了一个分配律。其余性质的证明留作练习。

表 10-5　布尔恒等式

恒等式	名称	恒等式	名称
$\overline{\overline{x}}=x$	双重补律	$x+yz=(x+y)(x+z)$ $x(y+z)=xy+xz$	分配律
$x+x=x$ $x\cdot x=x$	幂等律	$\overline{(xy)}=\overline{x}+\overline{y}$ $\overline{(x+y)}=\overline{x}\,\overline{y}$	德摩根律
$x+0=x$ $x\cdot 1=x$	同一律	$x+xy=x$ $x(x+y)=x$	吸收律
$x+1=1$ $x\cdot 0=0$	支配律	$x+\overline{x}=1$	单位元性质

（续）

恒等式	名称	恒等式	名称
$x+y=y+x$ $xy=yx$	交换律	$x\overline{x}=0$	零元性质
$x+(y+z)=(x+y)+z$ $x(yz)=(xy)x$	结合律		

例 6 证明分配律 $x(y+z)=xy+xz$ 是正确的。

解 表 10-6 表示此恒等式的验证。这个恒等式成立是因为此表的最后两列相同。■

表 10-6

x	y	z	$y+z$	xy	xz	$x(y+z)$	$xy+xz$
1	1	1	1	1	1	1	1
1	1	0	1	1	0	1	1
1	0	1	1	0	1	1	1
1	0	0	0	0	0	0	0
0	1	1	1	0	0	0	0
0	1	0	1	0	0	0	0
0	0	1	1	0	0	0	0
0	0	0	0	0	0	0	0

布尔代数恒等式可以被用来进一步证明其他的恒等式，如例 7 所示。

Extra Examples **例 7** 用表 10-5 所示的布尔代数的其他恒等式证明吸收律 $x(x+y)=x$（称为吸收律是因为将 $x+y$ 吸收进 x 而保持 x 不变。）

解 推导此恒等式的步骤及每步使用的定律如下：

$x(x+y)=(x+0)(x+y)$	布尔和的同一律
$=x+0\cdot y$	布尔和对布尔积的分配律
$=x+y\cdot 0$	布尔积的交换律
$=x+0$	布尔积的支配律
$=x$	布尔和的同一律 ■

10.1.4 对偶性

表 10-5 中的恒等式都是成对出现的(除了双重补律、单位元性质及零元性质)。为解释每一对恒等式中的两个式子的关系，我们使用“对偶”这个概念。一个布尔表达式的对偶可如下得到：交换布尔和与布尔积，且交换 0 与 1。

例 8 写出 $x(y+0)$ 和 $\overline{x}\cdot 1+(\overline{y}+z)$ 的对偶。

解 在这两个表达式中交换符号＋和·、0 和 1 就产生它们的对偶，这两个对偶分别是 $x+(y\cdot 1)$ 和 $(\overline{x}+0)(\overline{y}z)$。■

布尔表达式所表示的布尔函数 F 的对偶是由这个表达式的对偶所表示的函数，这个对偶函数记为 F^d，它不依赖于表示 F 的那个特定的布尔表达式。对于由布尔表达式表示的函数的恒等式，当取恒等式两边的函数的对偶时，等式仍然成立(原因参见练习 28)，此结果叫做**对偶性原理**，它对于获得新的恒等式十分有用。

例 9 通过取对偶的方法，由吸收律 $x(x+y)=x$ 构造一个恒等式。

解 取此恒等式两边的对偶，得到恒等式 $x+xy=x$，它也被称为吸收律，见表 10-5。■

10.1.5 布尔代数的抽象定义

虽然本节着重讨论布尔函数和表达式，但所有的结论都可以翻译成关于命题的结论，也可以翻译成关于集合的结论。因此，抽象地定义布尔代数十分有用。一旦一个特定的结构被证明是布尔代数，则所有关于布尔代数的一般结果都可应用于这个特定的结构。

布尔代数可以用多种方法来定义，最常用的方法是指明运算所必须满足的性质，如定义 1 所示。

定义 1 一个布尔代数是一个集合 B，它有两个二元运算 $\vee$ 和 $\wedge$，元素 0 和 1，以及一个一元运算 $^{-}$，且对 B 中的所有元素 x、y 和 z，下列性质成立：

$$\left.\begin{aligned} x \vee 0 &= x \\ x \wedge 1 &= x \end{aligned}\right\} \quad \text{同一律}$$

$$\left.\begin{aligned} x \vee \overline{x} &= 1 \\ x \wedge \overline{x} &= 0 \end{aligned}\right\} \quad \text{补律}$$

$$\left.\begin{aligned} (x \vee y) \vee z &= x \vee (y \vee z) \\ (x \wedge y) \wedge z &= x \wedge (y \wedge z) \end{aligned}\right\} \quad \text{结合律}$$

$$\left.\begin{aligned} x \vee y &= y \vee x \\ x \wedge y &= y \wedge x \end{aligned}\right\} \quad \text{交换律}$$

$$\left.\begin{aligned} x \vee (y \wedge z) &= (x \vee y) \wedge (x \vee z) \\ x \wedge (y \vee z) &= (x \wedge y) \vee (x \wedge z) \end{aligned}\right\} \quad \text{分配律}$$

使用定义 1 所给的定律可以证明许多其他的定律，例如幂等律、支配律等。（见练习 31～38。）

以前讨论过，$B=\{0, 1\}$ 连同 OR、AND 运算及“补”运算满足所有这些性质。有 n 个变元的所有命题构成的集合，连同 $\vee$ 和 $\wedge$ 算子、**F** 和 **T** 及“非”算子，也满足布尔代数的所有性质，这可以从 1.2 节中的表 1-13 看出来。类似地，一个全集 U 的所有子集构成的集合，连同并和交运算、空集和全集及集合的余算子是一个布尔代数，这可以从 1.5 节的表 1-21 中看出来。所以，为了建立关于布尔表达式、命题和集合的结果，我们只要证明关于抽象布尔代数的结果即可。

布尔代数也可以用第 7 章中所讨论的格的概念来定义。一个格 L 是一个偏序集，其每对元素 x、y 都有一个最小上界，记为 $\mathrm{lub}(x, y)$，也有一个最大下界，记为 $\mathrm{glb}(x, y)$。给定 L 的两个元素 x 和 y，我们可以定义 L 的两个运算 $\vee$ 和 $\wedge$ 如下：$x \vee y=\mathrm{lub}(x, y)$ 且 $x \wedge y=\mathrm{glb}(x, y)$。

要使一个格 L 成为定义 1 所指出的一个布尔代数，它必须还有两个性质。第一，它必须是可补的。为使一个格成为可补的，它必须有一个最小元 0 和一个最大元 1，且对格的每个元素 x，必须存在一个元素 $\overline{x}$，使得 $x \vee \overline{x}=1$ 且 $x \wedge \overline{x}=0$。第二，它必须是分配的。所谓“分配的”是指：对于 L 中的每个 x、y 和 z，$x \vee (y \wedge z)=(x \vee y) \wedge (x \vee z)$ 且 $x \wedge (y \vee z)=(x \wedge y) \vee (x \wedge z)$。证明可补分配格是布尔代数在本节末尾留作练习 39。

练习

1. 求下列表达式的值。

   a) $1 \cdot \overline{0}$　　b) $1+\overline{1}$　　c) $\overline{0} \cdot 0$　　d) $\overline{(1+0)}$

2. 求满足下列方程的布尔变元 x 的值(如果有的话)。

   a) $x \cdot 1=0$　　b) $x+x=0$　　c) $x \cdot 1=x$　　d) $x \cdot \overline{x}=1$

3. 用表来表示下列每个布尔函数的值。

a) $F(x, y, z)=\bar{x}y$ b) $F(x, y, z)=x+yz$

c) $F(x, y, z)=x\bar{y}+\overline{(xyz)}$ d) $F(x, y, z)=x(yz+\bar{y}\bar{z})$

4. 用表来表示下列每个布尔函数的值。

a) $F(x, y, z)=\bar{z}$ b) $F(x, y, z)=\bar{x}y+\bar{y}z$

c) $F(x, y, z)=x\bar{y}z+\overline{(xyz)}$ d) $F(x, y, z)=\bar{y}(xz+\overline{xz})$

5. 用 3 立方体 Q_3 表示练习 3 中的每个布尔函数，将函数值为 1 的 3 元组所对应的顶点画成黑圈。

6. 用 3 立方体 Q_3 表示练习 4 中的每个布尔函数，将函数值为 1 的 3 元组所对应的顶点画成黑圈。

7. 布尔变元 x 和 y 取什么值可满足 $xy=x+y$?

8. 有多少个 7 度布尔函数?

9. 用表 10-5 中的定律证明吸收律 $x+xy=x$。

☞10. 证明 $F(x, y, z)=xy+xz+yz$ 取值 1 当且仅当变元 x、y 和 z 中至少有两个取值 1。

11. 证明 $x\bar{y}+y\bar{z}+\bar{x}z=\bar{x}y+\bar{y}z+x\bar{z}$。

练习 12～21 处理由{0, 1}上的布尔和与布尔积所定义的布尔代数。对每一题，采用与例 6 中表 10-6 相似的形式进行验证。

12. 验证双重补律。

13. 验证幂等律。

14. 验证同一律。

15. 验证支配律。

16. 验证交换律。

17. 验证结合律。

18. 验证表 10-5 中的第一个分配律。

19. 验证德摩根律。

20. 验证单位元性质。

21. 验证零元性质。

布尔算子⊕的定义如下：$1\oplus1=0$，$1\oplus0=1$，$0\oplus1=1$，$0\oplus0=0$。此算子被称为"异或(XOR)"算子。

22. 化简下列表达式。

a) $x\oplus0$ b) $x\oplus1$ c) $x\oplus x$ d) $x\oplus\bar{x}$

23. 证明下列恒等式成立。

a) $x\oplus y=(x+y)\overline{(xy)}$ b) $x\oplus y=(x\bar{y})+(\bar{x}y)$

24. 证明 $x\oplus y=y\oplus x$。

25. 证明下列等式成立或不成立。

a) $x\oplus(y\oplus z)=(x\oplus y)\oplus z$ b) $x+(y\oplus z)=(x+y)\oplus(x+z)$ c) $x\oplus(y+z)=(x\oplus y)+(x\oplus z)$

26. 求下列布尔表达式的对偶。

a) $x+y$ b) $\bar{x}\bar{y}$ c) $xyz+\bar{x}\bar{y}\bar{z}$ d) $x\bar{z}+x\cdot0+\bar{x}\cdot1$

*27. 设 F 是一个布尔函数，它由一个含有变元 x_1, …, x_n 的布尔表达式表示。证明 $F^d(x_1, \cdots, x_n)=\overline{F(\overline{x_1}, \cdots, \overline{x_n})}$。

*28. 设 F 和 G 是由 n 个变元的布尔表达式表示的布尔函数，$F=G$。证明 $F^d=G^d$，其中 F^d 和 G^d 分别是由表示 F 和 G 的布尔表达式的对偶表示的布尔函数。(提示：使用练习 27 的结果。)

*29. 有多少个不同的布尔函数 $F(x, y, z)$ 使得：对布尔变元 x、y、z 的所有值，$F(\bar{x}, \bar{y}, \bar{z})=F(x, y, z)$。

*30. 有多少个不同的布尔函数 $F(x, y, z)$ 使得：对布尔变元 x、y、z 的所有值，$F(\bar{x}, y, z)=F(x, \bar{y}, z)=F(x, y, \bar{z})$。

在练习31～38中，用定义1中的定律来证明所述性质对每个布尔代数成立。

31. 在布尔代数中证明：幂等律 $x \vee x = x$ 和 $x \wedge x = x$ 对每个元素 x 成立。
32. 在布尔代数中证明：每个元素 x 都有唯一的一个补 $\overline{x}$ 使得 $x \vee \overline{x} = 1$ 且 $x \wedge \overline{x} = 0$。
33. 在布尔代数中证明：每个元素0的补是1，反之也成立。
34. 证明双重补律在布尔代数中成立，即对每个元素 x，$\overline{\overline{x}} = x$。
35. 证明德摩根律在布尔代数中成立，即对任意元素 x 和 y，$\overline{(x \vee y)} = \overline{x} \wedge \overline{y}$ 且 $\overline{(x \vee y)} = \overline{x} \vee \overline{y}$。
36. 证明模性质在布尔代数中成立，即证明：$x \wedge (y \vee (x \wedge z)) = (x \wedge y) \vee (x \wedge z)$ 且 $x \vee (y \wedge (x \vee z)) = (x \vee y) \wedge (x \vee z)$。
37. 在布尔代数中证明：如果 $x \vee y = 0$，则 $x = 0$ 且 $y = 0$；如果 $x \wedge y = 1$，则 $x = 1$ 且 $y = 1$。
38. 在布尔代数中证明一个恒等式的对偶还是一个恒等式，其中，恒等式的对偶是如下得到的：交换 $\wedge$ 和 $\vee$ 运算，还要交换元素0和1。
39. 证明一个可补的分配格是一个布尔代数。

10.2 布尔函数的表示

本节将研究布尔代数的两个重要问题。第一个是：给定一个布尔函数，怎样才能找到表示这个布尔函数的布尔表达式？这个问题将通过证明如下结论来解决：任何一个布尔函数都可由变元及其补的布尔积的布尔和表示。这个问题的答案还说明了任意布尔函数都可用三个布尔算子（·，+和⁻）表示。第二个是：有没有一个更小的算子集合可以用来表示所有的布尔函数？我们将通过证明下列结论来解决这个问题：所有的布尔函数都可仅用一个算子来表示。这两个问题在电路设计中都有特殊的重要性。

10.2.1 积之和展开式

下面用例子来说明寻找表示布尔函数的布尔表达式的一个重要方法。

例1 函数 $F(x, y, z)$ 和 $G(x, y, z)$ 如表10-7所示，求表示这两个函数的布尔表达式。

表 10-7

x	y	z	F	G
1	1	1	0	0
1	1	0	0	1
1	0	1	1	0
1	0	0	0	0
0	1	1	0	0
0	1	0	0	1
0	0	1	0	0
0	0	0	0	0

解 我们需要这样一个表达式来表示 F：当 $x = z = 1$ 且 $y = 0$ 时它的值为1，否则它的值为0。此表达式可取为 x、$\overline{y}$ 和 z 的布尔积，这个积 $x\overline{y}z$ 具有值1当且仅当 $x = \overline{y} = z = 1$，即当且仅当 $x = z = 1$ 且 $y = 0$。

为表示 G，我们需要一个表达式满足：当 $x = y = 1$ 且 $z = 0$ 时，或当 $x = z = 0$ 且 $y = 1$ 时，它的值为1。这样的表达式可以取为两个不同的布尔积的布尔和。布尔积 $xy\overline{z}$ 具有值1当且仅当 $x = y = 1$ 且 $z = 0$；类似地，布尔积 $\overline{x}y\overline{z}$ 具有值1当且仅当 $x = z = 0$ 且 $y = 1$。这两个布尔积的布尔和 $xy\overline{z} + \overline{x}y\overline{z}$ 就表示 G，因为它具有值1当且仅当 $x = y = 1$ 且 $z = 0$，或 $x = z = 0$ 且 $y = 1$。 ■

例1说明了一个过程，用这个过程可以构造布尔表达式来表示具有给定值的函数。如果变元值的一个组合使得函数值为1，则此组合确定了变元或其补的一个布尔积。

定义1 布尔变元或其补称为文字。布尔变元 $x_1, x_2, \cdots, x_n$ 的小项是一个布尔积 $y_1 y_2 \cdots y_n$，其中 $y_i = x_i$，或 $y_i = \overline{x_i}$。因此小项是 n 个文字的积，每个文字对应于一个变元。

一个小项对一个且只对一个变元值的组合取值1，更确切地，小项 $y_1 y_2 \cdots y_n$ 为1当且仅当每个 y_i 为1，当且仅当 $y_i = x_i$ 时 x_i 为1，$y_i = \overline{x_i}$ 时 x_i 为0。

例 2 求一个小项使得：当 $x_1=x_3=0$ 且 $x_2=x_4=x_5=1$ 时，它为 1；否则为 0。

解 小项 $\overline{x_1}x_2\overline{x_3}x_4x_5$ 具有正确的值集合。■

通过取不同小项的布尔和，就能构造布尔表达式，使其具有给定的值集合。特别地，小项的布尔和具有值 1 仅当和中的某个小项具有值 1 时才成立；对于变元值的其他组合，它具有值 0。因此，给定一个布尔函数，可以构造小项的布尔和使得：当此布尔函数具有值 1 时它的值为 1，当此布尔函数具有值 0 时它的值为 0。此布尔和中的小项与使得此函数值为 1 的值的组合相对应。表示布尔函数的小项的和称为此函数的*积之和展开式*或*析取范式*。

例 3 求函数 $F(x,y,z)=(x+y)\bar{z}$ 的积之和展开式。

Extra Examples

解 下面用两种方法求 $F(x,y,z)$ 的积之和展开式。第一种方法是用布尔恒等式将这个积展开然后化简。过程如下：

$$\begin{aligned}F(x,y,z)&=(x+y)\bar{z}\\&=x\bar{z}+y\bar{z} &&\text{分配律}\\&=x1\bar{z}+1y\bar{z} &&\text{同一律}\\&=x(y+\bar{y})\bar{z}+(x+\bar{x})y\bar{z} &&\text{单位元性质}\\&=xy\bar{z}+x\bar{y}\bar{z}+xy\bar{z}+\bar{x}y\bar{z} &&\text{分配律}\\&=xy\bar{z}+x\bar{y}\bar{z}+\bar{x}y\bar{z} &&\text{幂等律}\end{aligned}$$

构造积之和展开式的第二种方法是：对 x、y 和 z 所有可能的取值都计算出 F 的值，这些值见表 10-8。F 的积之和展开式是三个小项的布尔和，这三个小项对应于表 10-8 的三行，它们使此函数的值为 1。从而 $F(x,y,z)=xy\bar{z}+x\bar{y}\bar{z}+\bar{x}y\bar{z}$。■

表 10-8

x	y	z	$x+y$	$\bar{z}$	$(x+y)\bar{z}$
1	1	1	1	0	0
1	1	0	1	1	1
1	0	1	1	0	0
1	0	0	1	1	1
0	1	1	1	0	0
0	1	0	1	1	1
0	0	1	0	0	0
0	0	0	0	1	0

也可以通过取布尔和的布尔积来求一个布尔表达式，使其表示一个布尔函数，所得到的展开式称为这个函数的*合取范式*或*和之积展开式*，这个展开式可以通过求积之和展开式的对偶而得到。本节末尾的练习 10 描述了怎样直接求这样的展开式。

10.2.2 函数完全性

每个布尔函数都可以表示为小项的布尔和。每个小项都是布尔变元或其补的布尔积。这说明了每个布尔函数都可以用布尔运算·、+和¯来表示。因为每个布尔函数都可以由布尔运算表示，我们称集合{·，+，¯}是*函数完全的*。还有没有更小的函数完全运算集合呢？如果这三个运算中的某一个能够由其余两个表示，则就还有。用德摩根律可以做到这一点。使用等式

$$x+y=\overline{\bar{x}\,\bar{y}}$$

可以消去所有布尔和，此等式可如下得到：先对 10.1 节中表 10-5 的第二个德摩根律的两边求补，再应用双重补律。这意味着{·，¯}是函数完全的。类似地，使用等式

$$xy=\overline{\bar{x}+\bar{y}}$$

可以消去所有布尔积，此等式可如下得到：先对 10.1 节中表 10-5 的第一个德摩根律的两边求补，再应用双重补律。因此，{+，¯}是函数完全的。注意，{+，·}不是函数完全的，因为用

这两个运算不可能表示布尔函数 $F(x)=\overline{x}$(见练习 19)。

我们已经找到了一些含有两个运算的函数完全集合，还能不能找到更小的集合——即只含一个运算的集合——它仍然是函数完全运算集合呢？这样的集合是存在的。定义运算"|"或"*NAND*"如下：$1\mid 1=0$ 且 $1\mid 0=0\mid 1=0\mid 0=1$。定义运算"↓"或"*NOR*"如下：$1\downarrow 1=1\downarrow 0=0\downarrow 1=0$ 且 $0\downarrow 0=1$。集合{|}和{↓}都是函数完全的。因为{·,‾}是函数完全的，故要说明{|}是函数完全的，只要证明两个运算·和‾都可以由运算|表示，这由下面两式完成：

$$\overline{x}=x\mid x$$

$$xy=(x\mid y)\mid(x\mid y)$$

读者应当验证这些等式(见练习 14)。证明{↓}的函数完全性留给读者(见练习 15 和 16)。

练习

1. 求布尔变元 x、y、z 或其补的布尔积，使得它具有值为 1 当且仅当
   a) $x=y=0$，$z=1$　　b) $x=0$，$y=1$，$z=0$
   c) $x=0$，$y=z=1$　　d) $x=y=z=0$
2. 求下列布尔函数的积之和展开式。
   a) $F(x,y)=\overline{x}+y$　　b) $F(x,y)=x\overline{y}$
   c) $F(x,y)=1$　　d) $F(x,y)=\overline{y}$
3. 求下列布尔函数的积之和展开式。
   a) $F(x,y,z)=x+y+z$　　b) $F(x,y,z)=(x+z)y$
   c) $F(x,y,z)=x$　　d) $F(x,y,z)=x\overline{y}$
4. 求布尔函数 $F(x,y,z)$ 的积之和展开式，$F(x,y,z)$ 等于 1 当且仅当
   a) $x=0$　　b) $xy=0$　　c) $x+y=0$　　d) $xyz=0$
5. 求布尔函数 $F(w,x,y,z)$ 的积之和展开式，$F(w,x,y,z)$ 等于 1 当且仅当 w、x、y 和 z 中有奇数个变元的值为 1。
6. 求布尔函数 $F(x_1,x_2,x_3,x_4,x_5)$ 的积之和展开式，$F(x_1,x_2,x_3,x_4,x_5)$ 等于 1 当且仅当 x_1，x_2，x_3，x_4 和 x_5 中至少有三个变元的值为 1。

求表示布尔函数的布尔表达式的另一种方法是：构造文字之布尔和的布尔积。练习 7～11 涉及这种表示。

7. 求布尔和，它包含 x 或 $\overline{x}$、y 或 $\overline{y}$、z 或 $\overline{z}$，使得它具有值 1 当且仅当
   a) $x=y=1$，$z=0$　　b) $x=y=z=0$　　c) $x=z=0$，$y=1$
8. 求文字之布尔和的布尔积，使得它的值为 0 当且仅当 $x=y=1$ 且 $z=0$，或者 $x=z=0$ 且 $y=1$，或者 $x=y=z=0$。[提示：取在练习 7 的(a)、(b)、(c)部分求得的布尔和的布尔积。]
9. 设布尔和为 $y_1+y_2+\cdots+y_n$，其中：$y_i=x_i$ 或 $y_i=\overline{x}_i$。证明此布尔和对且只对变元值的一个组合取 0 值，这个组合为：若 $y_i=x_i$，则 $x_i=0$；若 $y_i=\overline{x}_i$，则 $x_i=1$。这样的布尔和叫做**大项**。
10. 证明布尔函数可以表示为大项的布尔积。此表示称为该函数的**和之积展开式**或**析取范式**(提示：对于使得函数值为 0 的每个变元值组合，此积都含有一个对应的大项。)
11. 求练习 3 中每个函数的和之积展开式。
12. 用运算·和‾表示下列布尔函数。
   a) $x+y+z$　　b) $x+\overline{y}(\overline{x}+z)$　　c) $\overline{(x+\overline{y})}$　　d) $\overline{x}(x+\overline{y}+\overline{z})$
13. 用运算+和‾表示练习 12 中的布尔函数。
14. 证明：
   a) $\overline{x}=x\mid x$　　b) $xy=(x\mid y)\mid(x\mid y)$　　c) $x+y=(x\mid x)\mid(y\mid y)$

15. 证明：

a) $\overline{x}=x\downarrow x$　　b) $xy=(x\downarrow x)\downarrow(y\downarrow y)$　　c) $x+y=(x\downarrow y)\downarrow(x\downarrow y)$

16. 利用练习 15 证明{$\downarrow$}是函数完全集。

17. 用运算 | 表示练习 3 中的布尔函数。

18. 用运算 $\downarrow$ 表示练习 3 中的布尔函数。

19. 证明运算集{+，·}不是函数完全的。

20. 下列运算集是否为函数完全的呢？

a) {+，⊕}　　b) {$^{-}$，⊕}　　c) {·，⊕}

10.3 逻辑门电路

10.3.1 引言

布尔代数被用来作为电子装置的电路模型，这种装置的输入和输出都可以认为是集合{0，1}中的元素。计算机或其他的电子装置就是由许多电路构成的，电路可以根据布尔代数的规则来设计，这些规则已经在 10.1 节和 10.2 节中讨论过。电路的基本元件是所谓的门，每种类型的门实现一种布尔运算。本节将定义几种类型的门，对这些门应用布尔代数的结果，就可以设计出电路来执行各种任务。在本章所讨论的电路中，输出都只与输入有关，而与电路的当前状态无关，换句话说，这些电路都没有存储能力，这样的电路叫做*组合电路*或*选通网络*。

我们将使用三种元件来构造组合电路，第一种是*反相器*，它以布尔值作为输入，并产生此布尔值的补作为输出。用来表示反相器的符号如图 10-2a 所示，进入元件的输入画在左边，离开元件的输出画在右边。

第二种元件是*或门*，其输入是两个或两个以上的布尔值，输出是这些值的布尔和。用来表示或门的符号如图 10-2b 所示，进入元件的输入画在左边，离开元件的输出画在右边。

第三种元件是*与门*，其输入是两个或两个以上的布尔值，输出是这些值的布尔积。用来表示与门的符号如图 10-2c 所示，进入元件的输入画在左边，离开元件的输出画在右边。

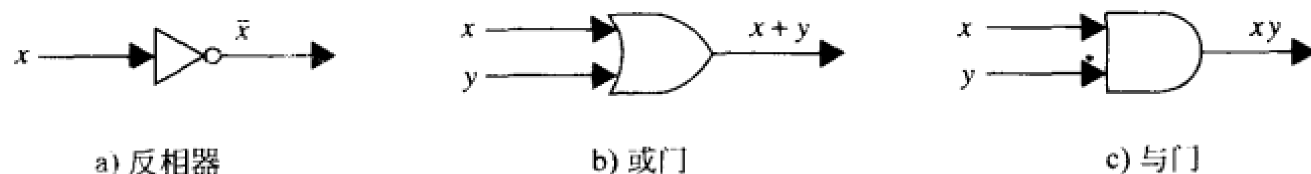


图 10-2　基本类型的门

与门和或门允许有多个输入，进入元件的输入都画在左边，离开元件的输出都画在右边。具有 n 个输入的与门和或门如图 10-3 所示。



图 10-3　具有 n 个输入的门

10.3.2 门的组合

使用反相器、或门和与门的组合可以构造组合电路。在构造电路的组合时，某些门可能有公共的输入。有两种方法可以描述公共输入，一种方法是：对每个输入，将使用这个输入的门画在

一个分支上；另一种方法是：对每个门，分别指出其输入。图 10-4 说明了这两种方法，其中的门有同样的输入值。注意，一个门的输出可能被作为另一个元件或更多元件的输入，如图 10-4 所示。图 10-4 中的两辐画描述了输出为 $xy+\bar{x}y$ 的电路。

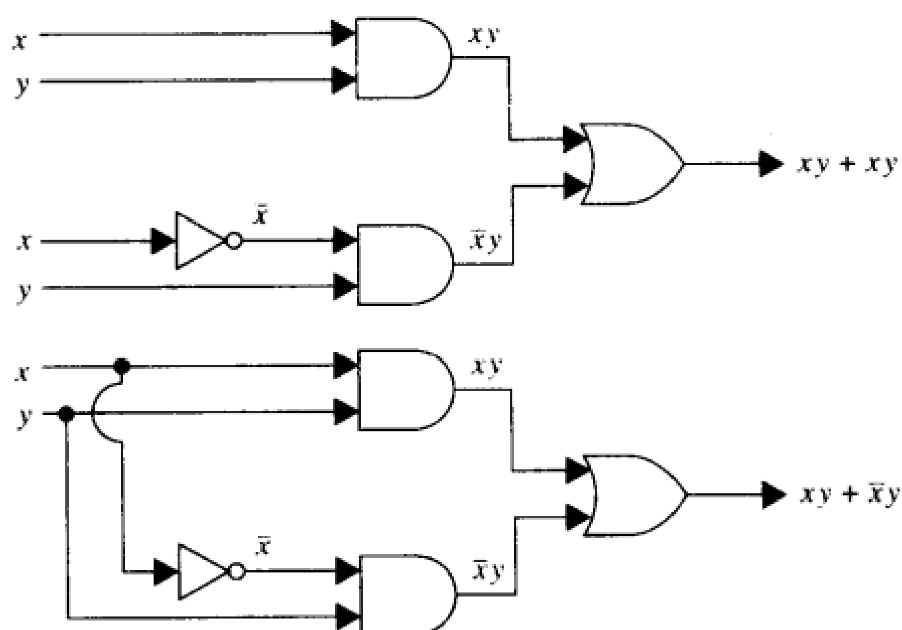


图 10-4 画相同电路的两种方法

例 1 构造产生下列输出的电路：a) $(x+y)\bar{x}$，b) $\bar{x}\,\overline{(y+\bar{z})}$，c) $(x+y+z)\bar{x}\,\bar{y}\,\bar{z}$。

解 产生这些输出的电路如图 10-5 所示。 ■

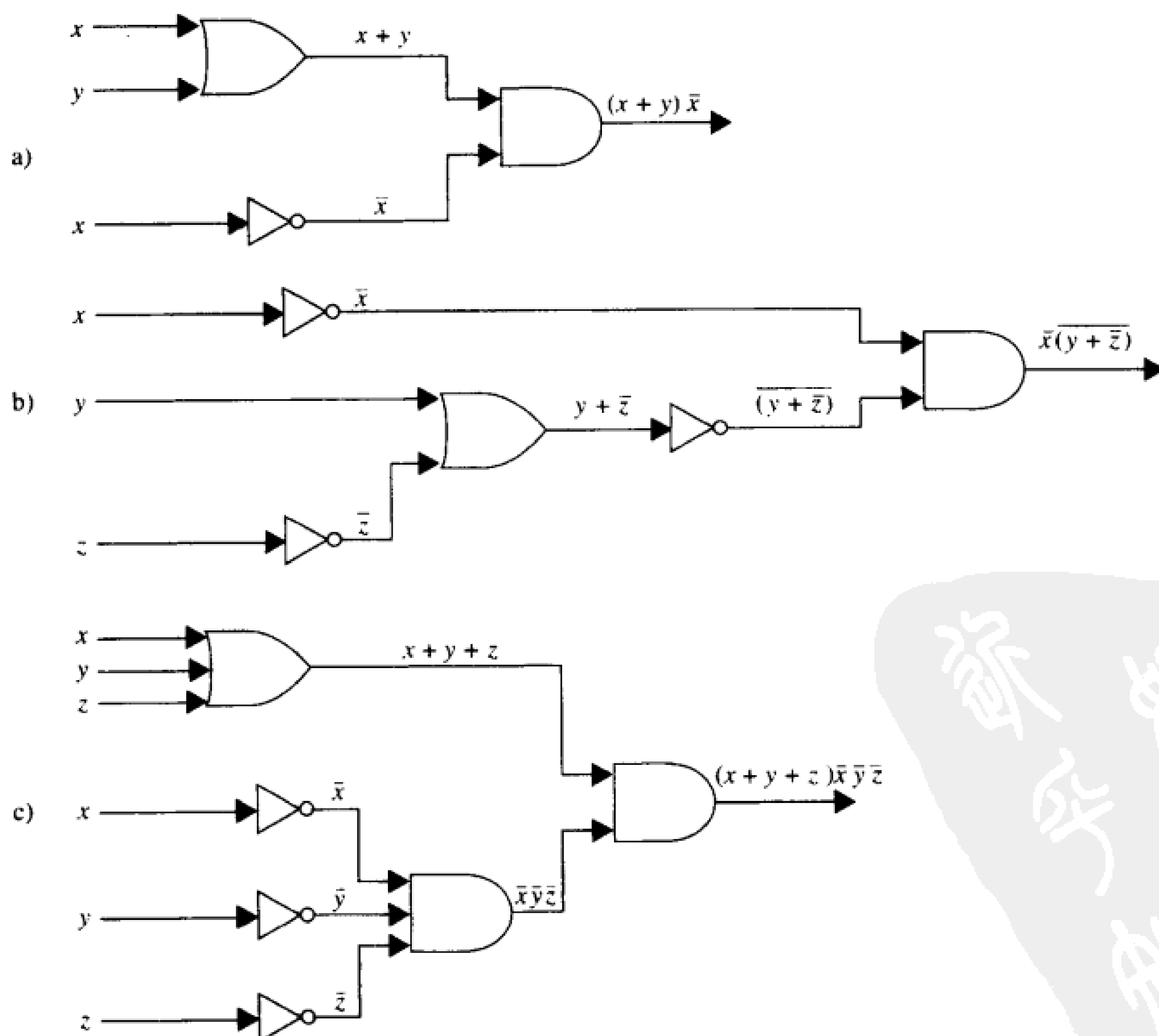


图 10-5 产生例 1 指出的输出的电路

10.3.3 电路的例子

下面给出一些具有实际功能的电路。

例 2 某个组织的一切事务都由一个三人委员会决定，每个委员对提出的建议可以投赞成票或反对票。一个建议如果得到了至少两张赞成票就获通过。设计一个电路，来判断建议是否获得通过。

Extra Examples **解** 如果第一个委员投赞成票，则令 $x=1$；如果这个委员投反对票，则令 $x=0$。如果第二个委员投赞成票，则令 $y=1$；如果这个委员投反对票，则令 $y=0$。如果第三个委员投赞成票，则令 $z=1$；如果这个委员投反对票，则令 $z=0$。必须设计一个电路使得对于输入 x、y 和 z，如果其中至少有两个为 1，则此电路产生输出 1。具有这样的输出值的一个布尔函数表示是 $xy+xz+yz$(见 10.1 节练习 10)。实现这个函数的电路如图 10-6 所示。 ■

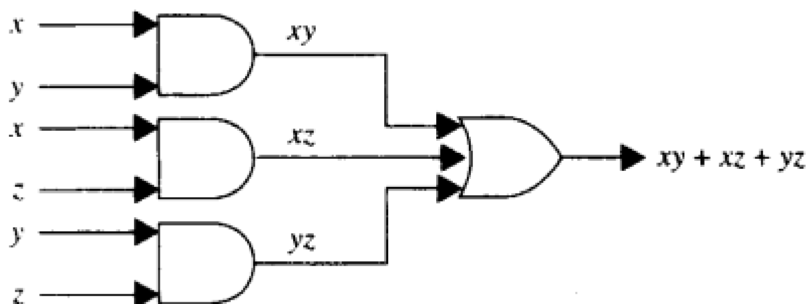


图 10-6 用于多数表决的电路

例 3 有时候灯具需要由多个开关来控制，因此有必要设计这样的电路：当灯是关闭时，敲击任何一个开关都可以打开此灯；反之，当灯是打开时，敲击任何一个开关都可以关闭此灯。在有两个开关或三个开关两种情形下，设计电路来完成这个任务。

解 首先设计使用两个开关的电路来控制灯具。当第一个开关关闭时，令 $x=1$；当它打开时，令 $x=0$。当第二个开关关闭时，令 $y=1$；当它打开时，令 $y=0$。当灯是打开时，令 $F(x, y)=1$；当灯是关闭时，令 $F(x, y)=0$。我们可以随意地假定：当两个开关都是关闭的时候，灯是打开的，即 $F(1, 1)=1$。这个假定决定了 F 的所有其他值：当两个开关中有一个是打开的时候，灯变灭了，故 $F(1, 0)=F(0, 1)=0$；当第二个开关也被打开的时候，灯又变亮了，故 $F(0, 0)=1$。表 10-9 列出了这些值。我们可以看到 $F(x, y)=xy+\overline{x}\,\overline{y}$。实现这个函数的电路如图 10-7 所示。

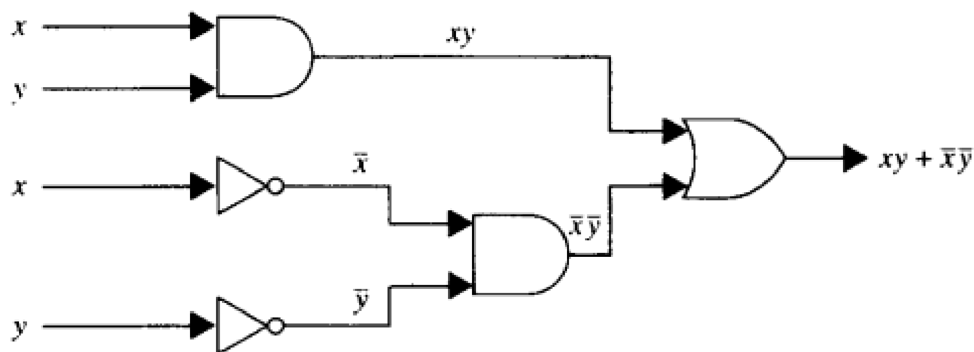


图 10-7 由两个开关控制的灯具的电路

表 10-9

x	y	$F(x, y)$
1	1	1
1	0	0
0	1	0
0	0	1

现在设计有三个开关的电路。设 x、y 和 z 是三个布尔变元，它们分别指出这三个开关是否是关闭的。当第一个开关处于关闭时，令 $x=1$；当它处于接通时，令 $x=0$。当第二个开关处于关闭时，令 $y=1$；当它处于接通时，令 $y=0$。当第三个开关处于关闭时，令 $z=1$；当它处于接通时，令 $z=0$。灯亮时，令 $F(x, y, z)=1$；灯不亮时，令 $F(x, y, z)=0$。当所有开关都关闭

时，我们可以随意地指定灯是亮的，即 $F(1, 1, 1)=1$，这确定了 F 的其他值。当一个开关被接通时，灯就变灭，故 $F(1, 1, 0)=F(1, 0, 1)=F(0, 1, 1)=0$。当又一个开关被接通时，灯又变亮了，故 $F(1, 0, 0)=F(0, 1, 0)=F(0, 0, 1)=1$。最后当三个开关都接通时，灯又变灭了，故 $F(0, 0, 0)=0$。这个函数的值如表 10-10 所示。

表 10-10

x	y	z	$F(x, y, z)$
1	1	1	1
1	1	0	0
1	0	1	0
1	0	0	1
0	1	1	0
0	1	0	1
0	0	1	1
0	0	0	0

函数 F 可以表示成积之和展开式：$F(x, y, z)=xyz+x\overline{y}\overline{z}+\overline{x}y\overline{z}+\overline{x}\overline{y}z$。实现这个函数的电路如图 10-8 所示。■

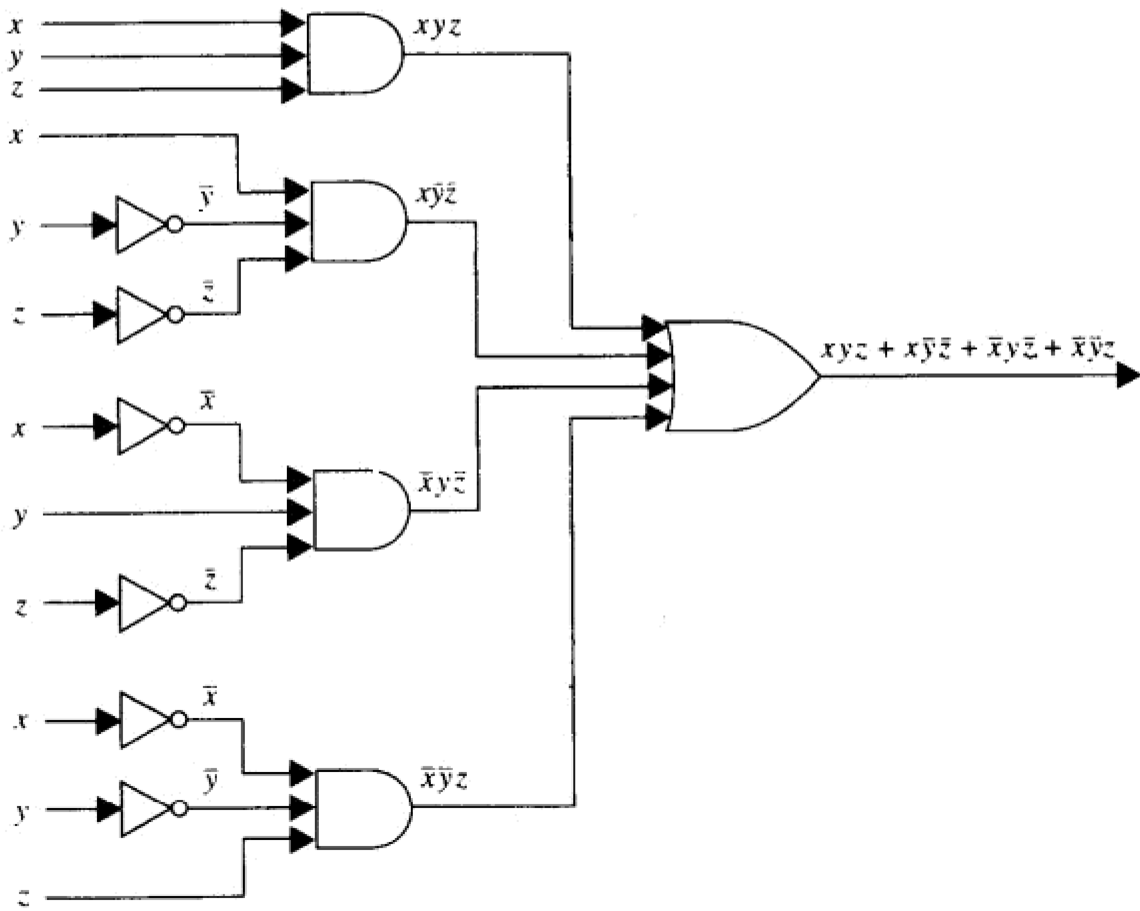


图 10-8 由三个开关控制的灯具的电路

10.3.4 加法器

下面说明如何用逻辑电路从两个正整数的二进制编码来执行加法。我们先构造一些分支电路，然后再从这些分支电路来构造加法电路。首先构造电路来计算 $x+y$，其中 x 和 y 是两个二进制数字。因为 x 和 y 的值为 0 或 1，此电路的输入就是 x 和 y。输出由两个二进制数字 s 和 c 构成，其中 s 和 c 分别是和位与进位。因为这种电路具有多个输出，故称为*多重输出电路*。又由于此电路只是将两个二进制数字相加，而没有考虑以前加法所产生的进位，所以这样电路称为*半加法器*。表 10-11 说明了半加法器的输入和输出。由此表可以看出 $c=xy$ 且 $s=x\overline{y}+\overline{x}y=(x+y)\overline{(xy)}$。图 10-9 所示的电路计算了 x 与 y 的和位 s 与进位 c。

当两个二进制数字与一个进位相加时，我们用*全加法器*来计算和位与进位。全加法器的输入是这两个二进制数字 x 和 y 以及此进位 c_i，输出是和位 s 与新的进位 c_{i+1}。全加法器的输入和输出如表 10-12 所示。

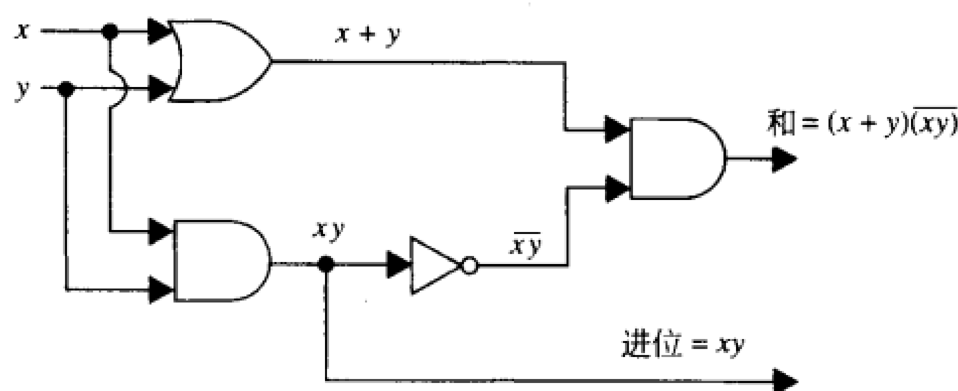


图 10-9 半加法器

表 10-11 半加法器的输入和输出

输入		输出	
x	y	s	c
1	1	0	1
1	0	1	0
0	1	1	0
0	0	0	0

表 10-12 全加法器的输入和输出

输入			输出	
x	y	c_i	s	c_{i+1}
1	1	1	1	1
1	1	0	0	1
1	0	1	0	1
1	0	0	1	0
0	1	1	0	1
0	1	0	1	0
0	0	1	1	0
0	0	0	0	0

全加法器的两个输出（和位 s 与进位 c_{i+1}）可分别由积之和展开式 $xyc_i+x\bar{y}\bar{c_i}+\bar{x}y\bar{c_i}+\bar{x}\bar{y}c_i$ 与 $xyc_i+xy\bar{c_i}+x\bar{y}c_i+\bar{x}yc_i$ 表示。但我们并不直接构造全加法器，而是使用半加法器来产生所需的输出。使用半加法器构造全加法器的方法如图 10-10 所示。

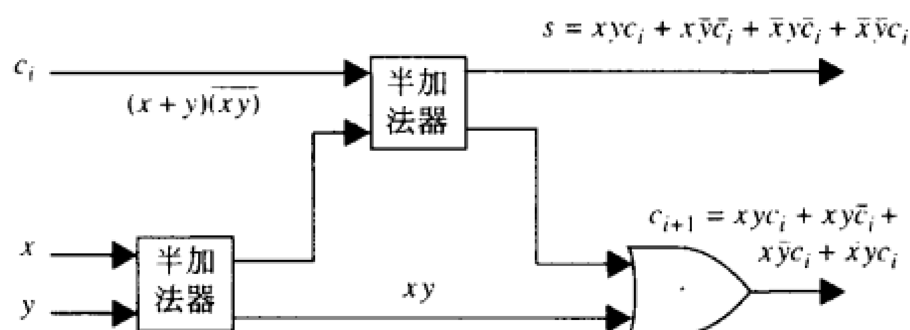


图 10-10 全加法器

最后，图 10-11 说明了怎样用全加法器和半加法器来计算两个 3 位二进制整数 $(x_2x_1x_0)_2$ 与 $(y_2y_1y_0)_2$ 之和 $(s_3s_2s_1s_0)_2$。注意，和中的最高位 s_3 是由进位 c_2 产生的。

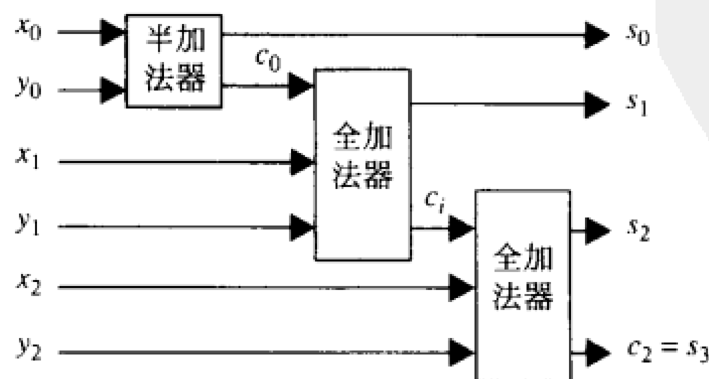
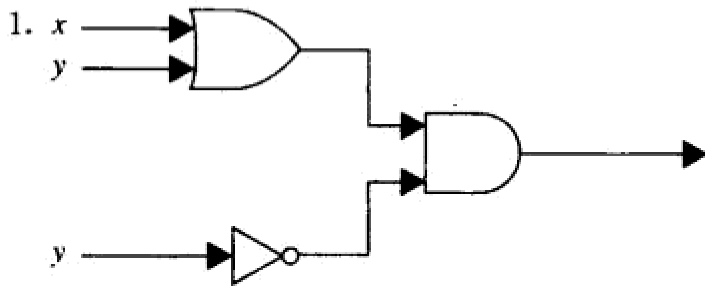
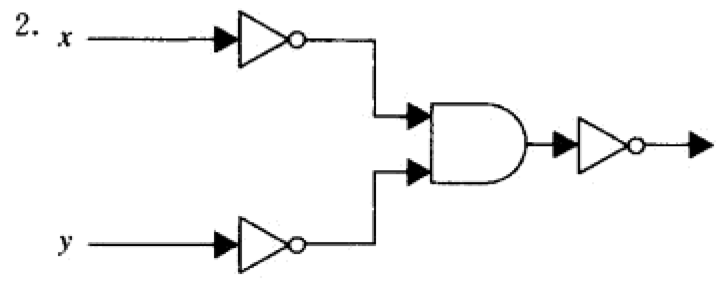


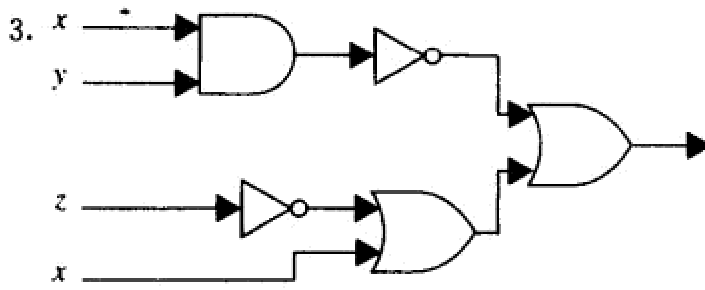
图 10-11 用全加法器和半加法器将两个 3 位整数相加

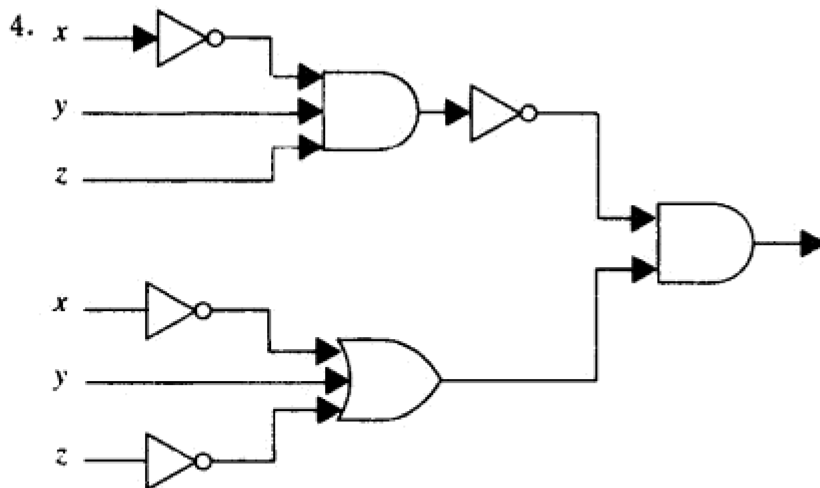
练习

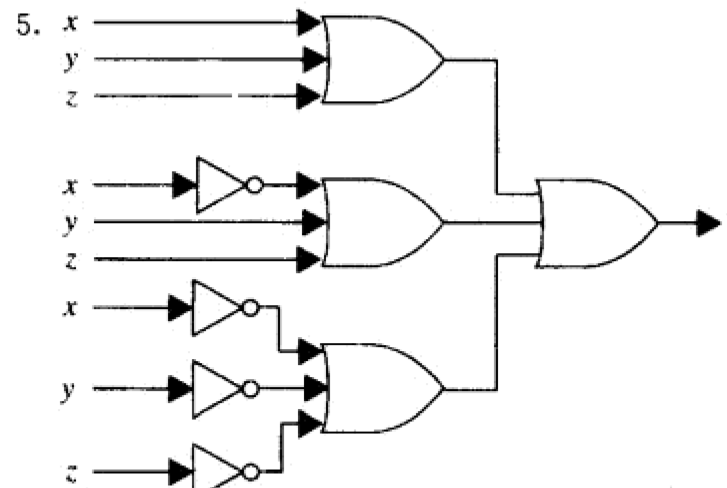
在练习 1～5 中，求所给电路的输出。

1. 

2. 

3. 

4. 

5. 

6. 用反相器、与门和或门构造产生下列输出的电路：

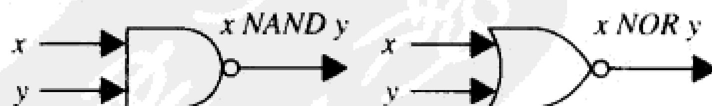
   a) $\overline{x}+y$　　b) $\overline{(x+y)}x$　　c) $xyz+\overline{x}\,\overline{y}\,\overline{z}$　　d) $\overline{(\overline{x}+z)(y+\overline{z})}$

7. 试设计一个电路来实现 5 个人的多数表决。
8. 试设计一个由四个开关控制的电灯混合控制器，使得当电灯在打开时，按动任意一个开关都可关闭它，或者当电灯在关闭时，按动任意一个开关都可打开它。
9. 证明可以使用全加法器和半加法器来计算两个 5 位二进制整数的和。
10. 一个半减法器的输入是两个二进制数字，输出是差和借位。试用与门、或门和反相器构造一个半减法器电路。
11. 一个全减法器的输入是两个二进制数字及一个借位，输出是差和借位。试用与门、或门和反相器构造一个全减法器电路。
12. 使用练习 10 和 11 中的电路计算两个 4 位二进制整数的差，其中第一个整数大于第二个整数。

*13. 构造一个电路来比较 2 位二进制整数 $(x_1x_0)_2$ 和 $(y_1y_0)_2$，使得当第一个整数大于第二个时，输出 1，否则输出 0。

*14. 构造一个计算 2 位二进制整数 $(x_1x_0)_2$ 与 $(y_1y_0)_2$ 之积的电路。此电路应该有四个输出位。

与非门和或非门也是电路中常用的两个门电路，如果使用这两个门来表示电路，就没有必要使用其他类型的门了。这两个门的记号如下：



*15. 使用与非门构造具有下列输出的电路：

a) $\bar{x}$　　b) $x+y$　　c) xy　　d) $x\oplus y$

*16. 使用或非门构造具有练习 15 中的输出的电路。

*17. 试用与非门构造半加法器。

*18. 试用或非门构造半加法器。

多路转接器是一种开关电路，它根据控制位的值将某组输入位输出。

19. 用与门、或门和反相器构造一个多路转接器，它的四个输入是二进制数字 x_0、x_1、x_2 和 x_3，控制位是 c_0 和 c_1。建造此电路使得 x_i 为输出，其中 i 是 2 位整数 $(c_1c_0)_2$ 的值。

组合电路的深度可如下定义：初始输入的深度为 0；如果一个门有 n 个输入，且其深度分别为 d_1，d_2，…，d_n，则它的输出的深度为 $\max(d_1, d_2, \cdots, d_n)+1$。这个值也定义为该门的深度。组合电路的深度为该电路中门的最大深度。

20. 求下列电路的深度：

a) 例 2 中构造的 3 人多数表决器。　　b) 例 3 中构造的 2 个开关控制的灯具。

c) 图 10-9 所示的半加法器。　　d) 图 10-10 所示的全加法器。

10.4 电路的极小化

10.4.1 引言

组合电路的有效性依赖于门的个数及安排。在组合电路的设计过程中，首先构造一个表，对于每种可能的输入值，此表说明对应的输出值。对于任一个电路，总可以用"积之和展开式"找到一组逻辑门来实现这个电路。但是，积之和展开式可能包含许多不必要的项。在一个积之和展开式中，若其中的一些项只在一个变元处不一样，即在某个项中此变元本身出现，而在另一个项中此变元的补出现，则这些项可以合成。例如，考虑这样的电路，它输出 1 当且仅当 $x=y=z=1$，或 $x=z=1$ 且 $y=0$。此电路的积之和展开式为 $xyz+x\bar{y}z$，在这个展开式的两个积中，只有一个变元以不同的形式出现，即 y。它们可以如下合并：

$$\begin{aligned} xyz+x\bar{y}z &= (y+\bar{y})(xz) \\ &= 1\cdot(xz) \\ &= xz \end{aligned}$$

这样，xz 也是一个表示这个电路的布尔表达式，但包含更少的算子。图 10-12 说明了这个电路的两个不同实现，第二个电路只使用一个门，但第一个却使用了三个门和一个反相器。

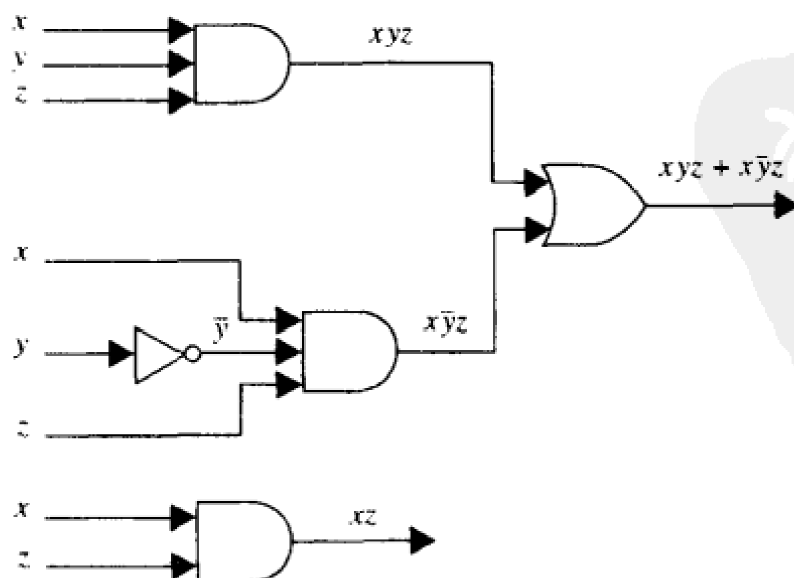


图 10-12　具有相同输出的两个电路

这个例子说明，在一个电路的积之和展开式中，将一些项合并会导出这个电路的更简单的表达式。下面将描述化简积之和展开式的两个过程。

这两个过程的目的都是为了产生布尔积的布尔和，使其所包含的文字之积的个数最少，从而使得这些积中在表示布尔函数的积之和中所包含的文字个数最少。寻求这种积之和称为**布尔函数的最小化**。最小化布尔函数可以为这个函数构造一条电路，这条电路在最小化布尔表达式的所有电路中，用最少的门和在电路中对 AND 门和 OR 门最少的输入。

直到 20 世纪 60 年代，逻辑门都是单独的组件。为了降低成本，采用最少的门得到期望的结果是非常重要的。在 20 世纪 60 年代中期，集成电路技术的发展使得将多个门组合到一个芯片成为可能。即使现在可以用非常低的成本对许多芯片构建非常复杂的集成电路，布尔函数的最小化仍然十分重要。减少芯片上门的数量带来更可靠的电路，并降低芯片的生产成本。同时，最小化还可以在同一芯片上设计更合适的电路。而且，最小化减少了电路中对门的输入。这减少了用电路计算出结果所用的时间，使对门的输入为有限的，因为在构建逻辑门电路采用了特殊的技术。

在第一个过程中，我们介绍 20 世纪 50 年代用手工最小化电路的、著名的卡诺图(或 K 图)。K 图在最多 6 个变量的最小化电路中非常有用，尽管对于 5、6 个变量来说已经变得相当复杂。第二个过程将介绍 20 世纪 60 年代发明的奎因一莫可拉斯基方法。这自动最小化组合电路的过程，由计算机程序实现。

遗憾的是，最小化有许多变量的布尔函数需要深入的计算。可以证明这个问题是 NP 完全问题(见 2. 3 节和[Ka93])，因此，最小化布尔电路的多项式时间算法也不太可能存在。奎因一莫可拉斯基方法具有指数复杂度。实际上，只能用于文字数量不超过 10 的情况。因为，20 世纪 70 年代在最小化组合电路方面开发了大量的新算法(见[Ha93]和[Ka93])。但是，还没有最好的算法，只能使不超过 25 个变量的电路最小化。同样，可以用启发式(或凭经验的)方法进行大量简化，但不一定是最小的有较大量变量的布尔表达式。

10.4.2 卡诺图

对于表示电路的一个布尔表达式，为了减少其中项的个数，有必要去发现可以合并的项。如果布尔函数所包含的变元相对较少，可以用一种图形法来发现能被合并的项，此法称为卡诺图，它是由 Maurice Karnaugh⊖ 在 1953 年发现的，他的方法建立在更早的 E. W. Veitch 的工作基础上(Veitch 的方法通常只适用于六个以下变元的函数)。卡诺图给出了一种化简积之和展开式的可视化方法，但此法不适用于将化简过程机械化。下面首先说明怎么用卡诺图来化简包含两个变元的布尔函数的展开式，然后说明如何用卡诺图来化简包含三个变元及四个变元的布尔函数，最后，我们将介绍卡诺图的扩展概念，可用以化简包含四个以上变元的布尔函数。

在具有两个变元 x 和 y 的布尔函数的和之积展开式中，有四种可能的小项。具有这两个变元的布尔函数的卡诺图由四个方格组成，如果一个小项在此展开式中出现，则表示这个小项的方格就被放置 1。如果一些方格所表示的小项只在一个处不一样，则称这些方格是**相邻的**。例如，表示$\bar{x}y$的方格与表示 xy 的方格及表示$\bar{x}\bar{y}$的方格都相邻。四个方格及其表示的项如图 10-13 所示。

	y	$\bar{y}$
x	xy	$x\bar{y}$
$\bar{x}$	$\bar{x}y$	$\bar{x}\bar{y}$

图 10-13 两个变元的卡诺图

⊖ **莫里斯·卡诺(Maurice Karnaugh，1924 年生)** 卡诺生于纽约，从纽约城市学院获得学士学位，从耶鲁大学获得硕士和博士学位。从 1952 年至 1966 年，卡诺一直在贝尔实验室从事技术工作，1966～1970 年，他任 AT&T 公司 FSD 分部的研发主管。1970 年，卡诺加入 IBM 成为研究员。卡诺对数字技术在计算和远程通信领域的应用做出了重要的贡献。他当前的研究兴趣包括计算机中基于知识的系统和启发式搜索方法。

例 1 找出下列各式的卡诺图：

(a)$xy+\bar{x}y$ (b)$x\bar{y}+\bar{x}y$ (c)$x\bar{y}+\bar{x}y+\bar{x}\bar{y}$

解 当一个方格所表示的小项在和之积展开式中出现时，我们就在这个方格中放置一个 1。三个卡诺图如图 10-14 所示。■

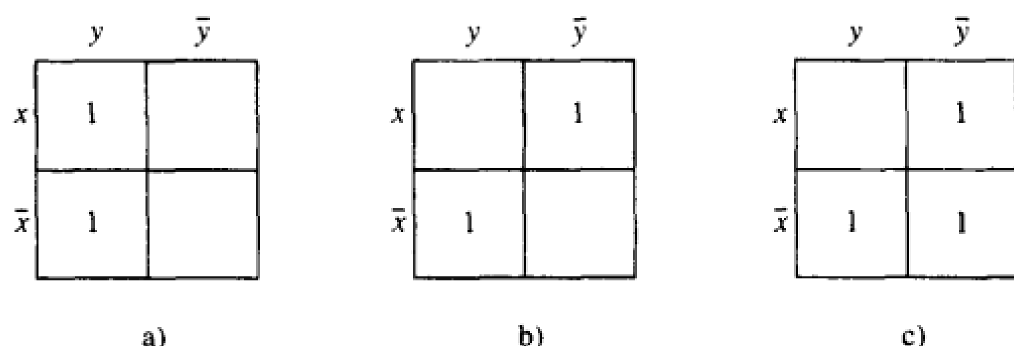


图 10-14 例 1 中积之和展开式的卡诺图

我们可以从卡诺图中识别出能够合并的小项。在卡诺图中，一旦有两个方格是相邻的，则由这两个方格所表示的小项就可被合并成一个积，且此积只涉及其中的一个变元。例如，$x\bar{y}$ 和 $\bar{x}\bar{y}$ 是由两个相邻的方格表示的，它们可以合并成 $\bar{y}$，因为 $x\bar{y}+\bar{x}\bar{y}=(x+\bar{x})\bar{y}=\bar{y}$。而且，如果所有四个方格都是 1，则四个小项可以合并成一个项，即布尔表达式 1，它不涉及任何变元。在卡诺图中，如果一些小项能够合并，则在卡诺图中，我们将表示这些小项的方格所组成的块用圆圈圈起来，然后找出对应的积之和，其目的是找出可能最大的块，以及以最少的块来覆盖所有的 1，在此过程中，首先使用最大的块，并总是使用可能最大的块。

例 2 化简例 1 中的积之和展开式。

解 用这些展开式的卡诺图对小项进行分组的方式如图 10-15 所示。这些积之和式的极小展开式是(a)y、(b)$x\bar{y}+\bar{x}y$ 和(c)$\bar{x}+\bar{y}$。■

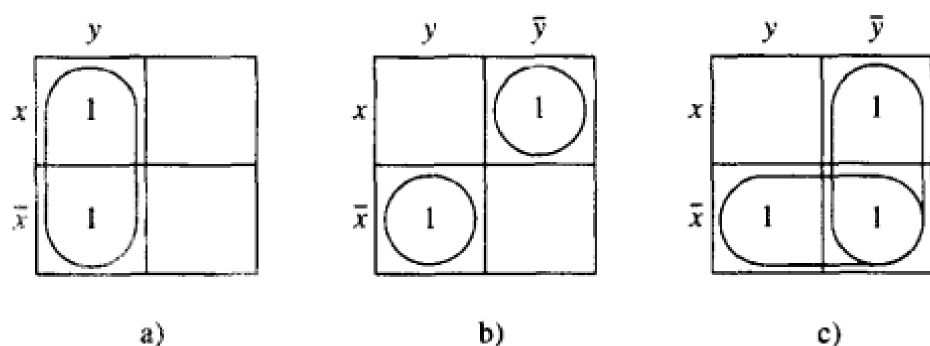


图 10-15 例 1 中积之和展开式的化简

三个变元的卡诺图是被分成八个方格的矩形，这些方格代表由三个变元组成的八个可能的小项。两个方格称为是相邻的，如果它们表示的小项只在一个文字处不一样。一种画三个变元卡诺图的方法如图 10-16a 所示，这个卡诺图可以被认为是贴在圆柱体的表面上，如图 10-16b 所示。在这个圆柱体的表面上，两个方格有公共边界当且仅当它们是相邻的。

为了化简三个变元的积之和展开式，我们用卡诺图来识别由可以合并的小项组成的块。两个相邻方格组成的块代表了一对小项，它们可以合并成两个文字的积，2×2 和 4×1 方格组成的块代表可以合并成一个文字的小项，全部八个方格组成的块代表函数 1，它不是任何文字的积。1×2、2×1、2×2、4×1 和 4×2 块及其代表的积如图 10-17 所示。

对应于卡诺图中全是 1 的块的文字之积被称为函数的**隐含数**。如果这个全 1 的块不包含在一个更大的由 1 组成的块中(且这个更大的块表示更少文字的积)，则称它为**素隐含数**。

我们的目的是在图中标出最大可能块，然后用最大块优先法则以最少的块覆盖所有的 1。最

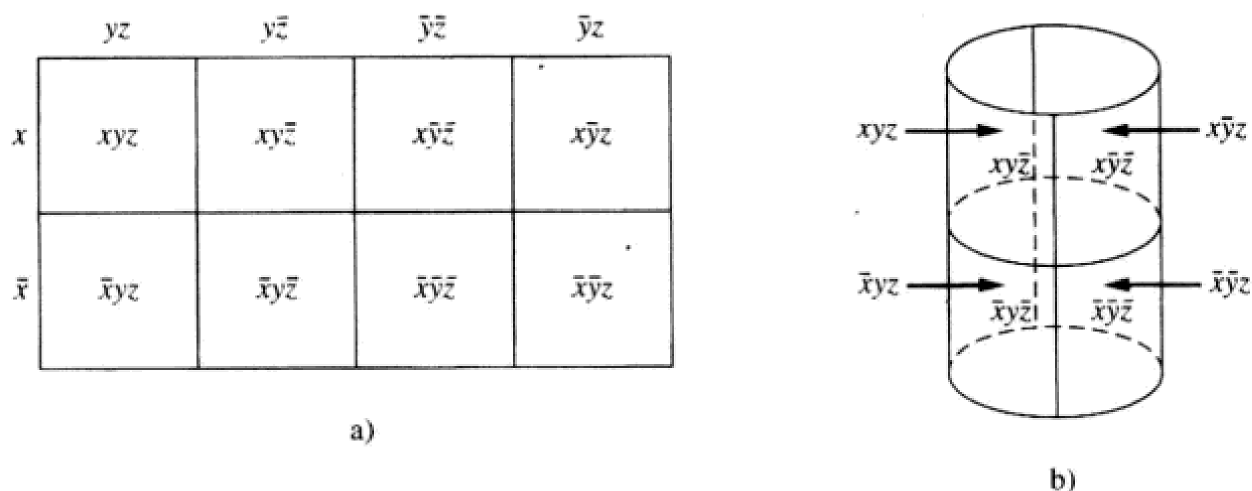


图 10-16　三个变元的卡诺图

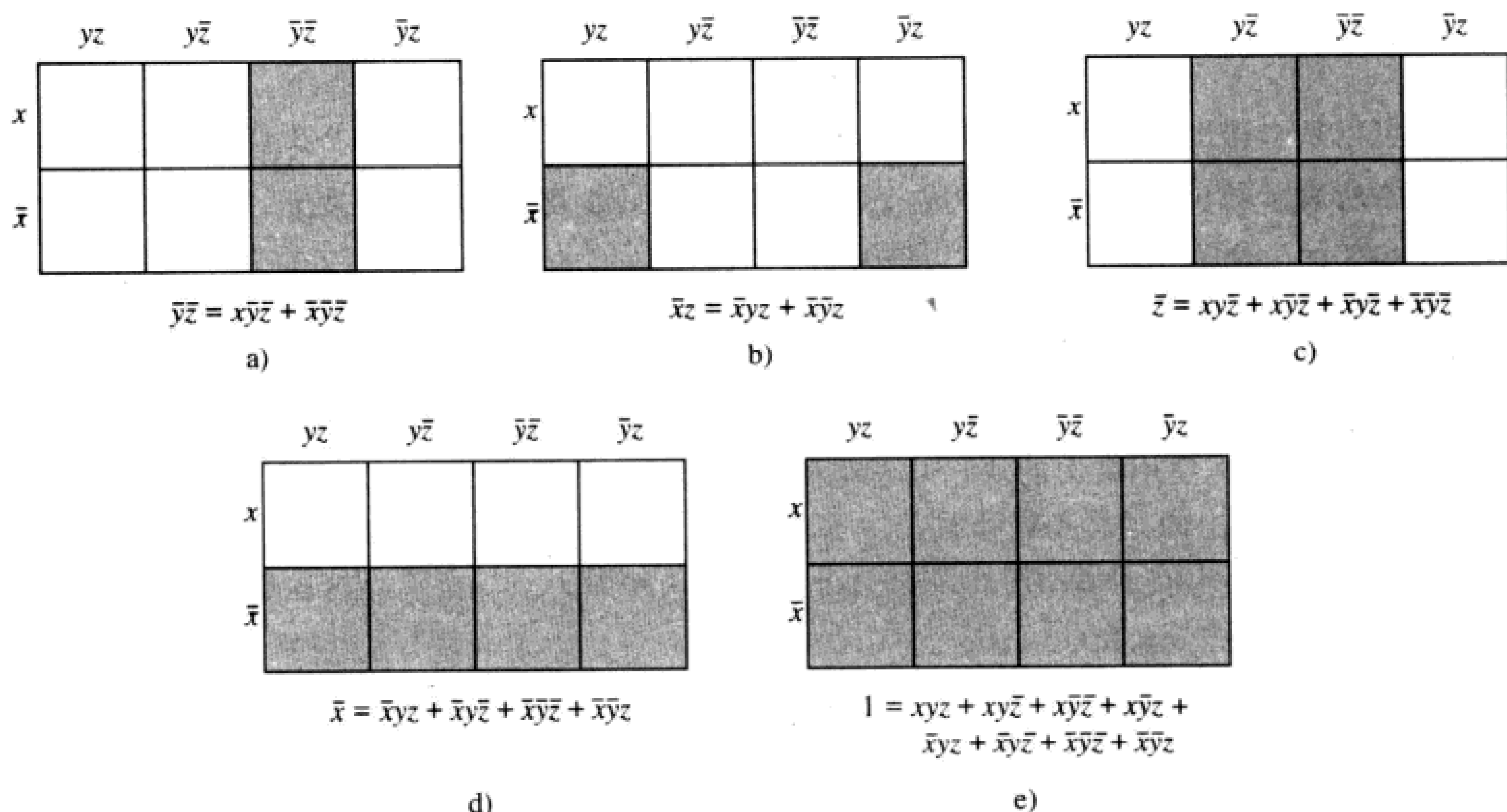


图 10-17　三个变元的卡诺图中的块

大可能的块总是会被选取，但如果卡诺图中只有一个块覆盖一个 1，则必须选取它，这样的块称为本原素隐含数。通过使用素隐含数对应的块来覆盖图中所有的 1，便可以用素隐含数之和来表达积之和。注意，以最少的块覆盖所有的 1 可能有不止一种方法。

例 3 示出了三变元卡诺图的使用。

例 3　用卡诺图化简下列积之和展开式：

(a)$xy\bar{z}+x\bar{y}\bar{z}+\bar{x}yz+\bar{x}\bar{y}\bar{z}$

(b)$x\bar{y}z+x\bar{y}\bar{z}+\bar{x}yz+\bar{x}\bar{y}z+\bar{x}\bar{y}\bar{z}$

(c)$xyz+xy\bar{z}+x\bar{y}z+x\bar{y}\bar{z}+\bar{x}yz+\bar{x}\bar{y}z+\bar{x}\bar{y}\bar{z}$

(d)$xy\bar{z}+x\bar{y}\bar{z}+\bar{x}\bar{y}z+\bar{x}\bar{y}\bar{z}$

解　这些积之和展开式的卡诺图如图 10-18 所示。块的分组表明：布尔积之布尔和展开式为(a)$x\bar{z}+\bar{y}\bar{z}+\bar{x}yz$、(b)$\bar{y}+\bar{x}z$、(c)$x+\bar{y}+z$ 和(d)$x\bar{z}+\bar{x}\bar{y}$。在(d)中注意，素隐含数 $x\bar{z}$ 和$\bar{x}\bar{y}$是本原素隐含数，但素隐含数$\bar{y}\bar{z}$则不是本原的，因为它覆盖的方格被其他两个素隐含数覆盖了。■

四个变元的卡诺图是被分成 16 个方格的正方形，这些方格代表由四个变元组成的 16 个可能的小项。一种画四个变元卡诺图的方法如图 10-19 所示。

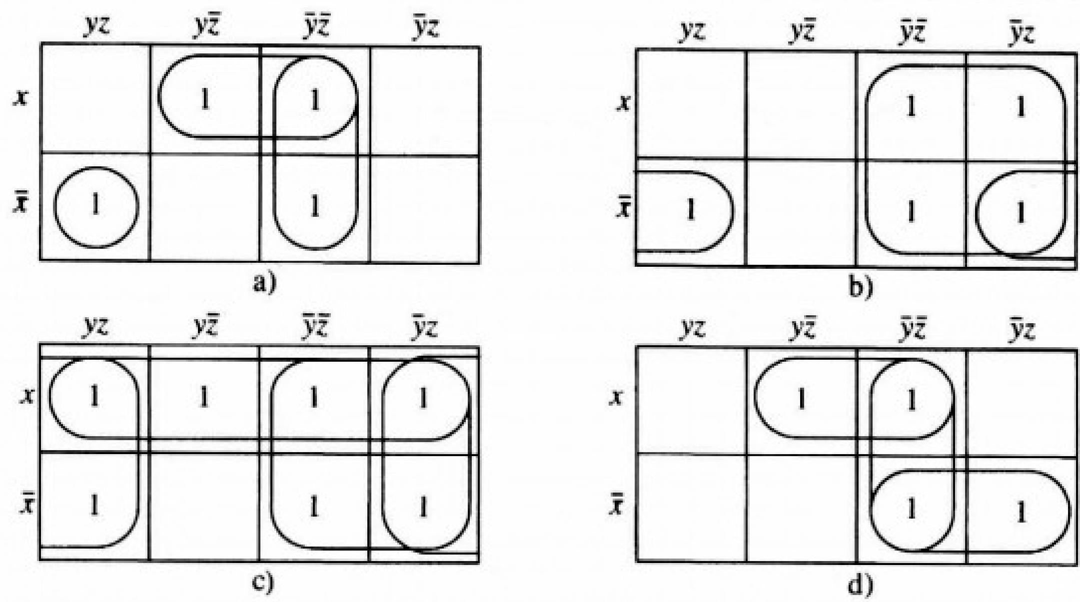
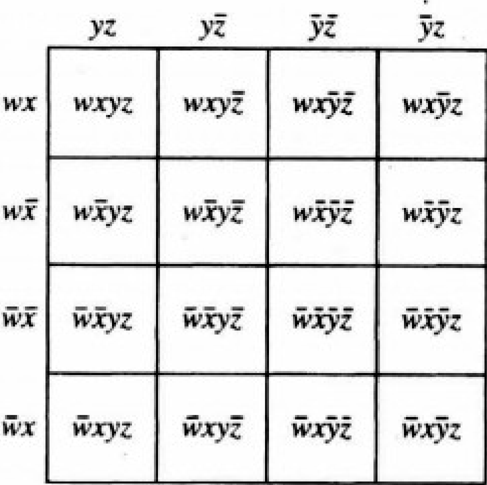


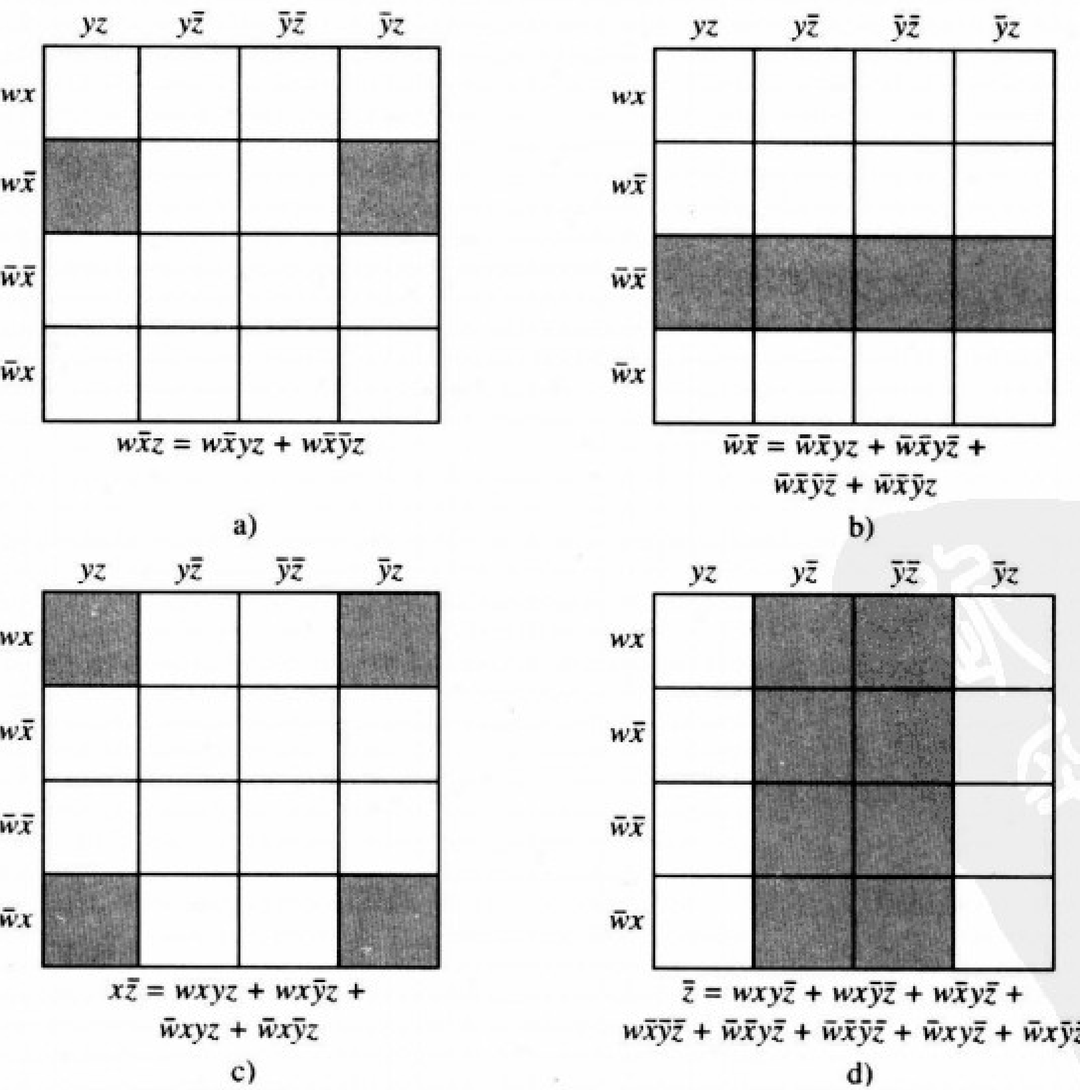
图 10-18 三变元卡诺图的使用

两个方格是相邻的当且仅当它们表示的小项只在一个文字处不一样。因而，每个方格都和另外四个方格相邻。四个变元的积之和展开式的卡诺图可以被认为是贴在圆环面上，因而相邻的方格具有公共的边界(见练习 28)。四个变元的积之和展开式的化简也是通过识别一些块来实现的，这些块可能由 2、4、8 或 16 个方格组成，它们代表的小项必须可以合并，且每个表示小项的方格都必须被用来产生更少个文字的积，或者包含在展开式中。在图 10-20 中，我们给出了一些块，这些块表示三个文字的积、两个文字的积或单个文字。



	yz	$y\bar{z}$	$\bar{y}\bar{z}$	$\bar{y}z$
wx	$wxyz$	$wxy\bar{z}$	$wx\bar{y}\bar{z}$	$wx\bar{y}z$
$w\bar{x}$	$w\bar{x}yz$	$w\bar{x}y\bar{z}$	$w\bar{x}\bar{y}\bar{z}$	$w\bar{x}\bar{y}z$
$\bar{w}\bar{x}$	$\bar{w}\bar{x}yz$	$\bar{w}\bar{x}y\bar{z}$	$\bar{w}\bar{x}\bar{y}\bar{z}$	$\bar{w}\bar{x}\bar{y}z$
$\bar{w}x$	$\bar{w}xyz$	$\bar{w}xy\bar{z}$	$\bar{w}x\bar{y}\bar{z}$	$\bar{w}x\bar{y}z$

图 10-19 四个变元的卡诺图



$w\bar{x}z = w\bar{x}yz + w\bar{x}\bar{y}z$

a)

$\bar{w}\bar{x} = \bar{w}\bar{x}yz + \bar{w}\bar{x}y\bar{z} + \bar{w}\bar{x}\bar{y}\bar{z} + \bar{w}\bar{x}\bar{y}z$

b)

$x\bar{z} = wxyz + wx\bar{y}z + \bar{w}xyz + \bar{w}x\bar{y}z$

c)

$\bar{z} = wxy\bar{z} + wx\bar{y}\bar{z} + w\bar{x}y\bar{z} + w\bar{x}\bar{y}\bar{z} + \bar{w}\bar{x}y\bar{z} + \bar{w}\bar{x}\bar{y}\bar{z} + \bar{w}xy\bar{z} + \bar{w}x\bar{y}\bar{z}$

d)

图 10-20 四变元卡诺图中的块

就像两个或三个变元的卡诺图的作用一样，我们的目的也是在图中标出 1 构成的最大块(对应于素隐含数)，然后用最大块优先法则以最少的块覆盖所有的 1，我们也总是使用可能最大的块。例 4 说明了四变元卡诺图的使用。

例 4 用卡诺图化简下列积之和展开式：

(a)$wxyz+wxy\bar{z}+wx\bar{y}\bar{z}+w\bar{x}yz+w\bar{x}\bar{y}z+w\bar{x}\bar{y}\bar{z}+\bar{w}x\bar{y}z+\bar{w}\bar{x}yz+\bar{w}\bar{x}y\bar{z}$

(b)$wx\bar{y}\bar{z}+w\bar{x}yz+w\bar{x}y\bar{z}+w\bar{x}\bar{y}\bar{z}+\bar{w}x\bar{y}\bar{z}+\bar{w}\bar{x}y\bar{z}+\bar{w}\bar{x}\bar{y}\bar{z}$

(c)$wxy\bar{z}+wx\bar{y}\bar{z}+w\bar{x}yz+w\bar{x}y\bar{z}+w\bar{x}\bar{y}\bar{z}+\bar{w}xyz+\bar{w}xy\bar{z}+\bar{w}x\bar{y}\bar{z}+\bar{w}x\bar{y}z+\bar{w}\bar{x}y\bar{z}+\bar{w}\bar{x}\bar{y}\bar{z}$

解 这些展开式的卡诺图如图 10-21 所示。用所示的块可导出如下的积之和：(a)$wyz+wx\bar{z}+w\bar{x}\bar{y}+\bar{w}\bar{x}y+\bar{w}x\bar{y}z$、(b)$\bar{y}\bar{z}+w\bar{x}y+\bar{x}\bar{z}$和(c)$\bar{z}+\bar{w}x+w\bar{x}y$。读者应该确定，是否可能选择其他的块，它们导致这些函数的不同积之和。 ■

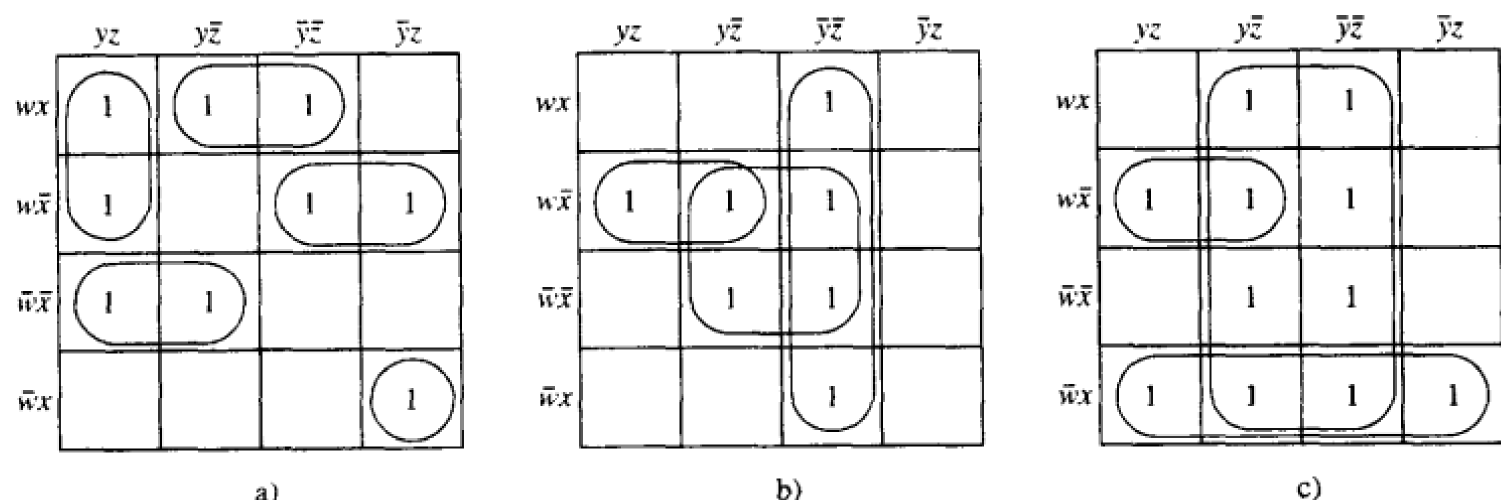


图 10-21 四变元卡诺图的使用

卡诺图可以实际用于化简五变元或六变元的布尔函数，但对更多变元的函数就很少使用卡诺图了，因为它们会非常复杂。然而，卡诺图中用到的概念在更新的算法中起着重要的作用。而且，掌握这些概念有助于理解这些算法及实现算法的计算机辅助设计(CAD)程序。在介绍这些概念时，会用到前面化简三变元、四变元布尔函数的内容。

用于化简两变元、三变元和四变元布尔函数的卡诺图分别是用 2×2、2×4 和 4×4 的矩形构建的。而且，在这些情形中，顶行和底行、最左列和最右列中的相应方格是相邻的，因为它们表示的小项只在一个文字处不同。我们可以用类似的方法构造四变元以上函数的卡诺图。我们使用包含 $2^{\lfloor n/2\rfloor}$行和 $2^{\lceil n/2\rceil}$列的矩形(这些卡诺图包含 2^n 个方格，因为$\lceil n/2\rceil+\lfloor n/2\rfloor=n$)。要构造具有这种相邻性的卡诺图，可以用格雷码来安排行和列(见 8.5 节)，其中通过指明 1 对应于变量的出现和 0 对应于变量的补的出现，可以将位串和积关联起来。例如，在一个 10 元卡诺图中，格雷码 01110 标记的行对应于积$\bar{x}_1x_2x_3x_4\bar{x}_5$。

例 5 用于化简四变元布尔函数的卡诺图有两行两列。行和列均用格雷码 11，10，00，01 来安排。行分别表示积 wx，$w\bar{x}$，$\bar{w}\bar{x}$和$\bar{w}x$。列分别对应积 yz，$y\bar{z}$，$\bar{y}\bar{z}$和$\bar{y}z$。使用格雷码并且第一行和最末行、第一列和最末列的方格相邻，我们确保了只在一个变元上不同的小项总是相邻的。 ■

例 6 要化简五变元的布尔函数，我们使用 $2^3=8$ 列和 $2^2=4$ 行的卡诺图。使用格雷码 11，10，00，01 标记四行，分别对应于 x_1x_2，$x_1\bar{x}_2$，$\bar{x}_1\bar{x}_2$ 和$\bar{x}_1x_2$。使用格雷码 111，110，100，101，001，000，010，011 标记八个列，分别对应项 $x_3x_4x_5$，$x_3x_4\bar{x}_5$，$x_3\bar{x}_4\bar{x}_5$，$x_3\bar{x}_4x_5$，$\bar{x}_3\bar{x}_4x_5$，$\bar{x}_3\bar{x}_4\bar{x}_5$，$\bar{x}_3x_4\bar{x}_5$ 和$\bar{x}_3x_4x_5$。使用格雷码标记行列确保了相邻方格表示的小项只在

一个变元上不同。然而，要确保所有只在一个变元上不同的表示积的方格是相邻的，要考虑顶行和底行的方格是相邻的，以及第一列和第八列的方格的相邻性，以及第一和第四列，第二和第七列，第三和第六列，第五和第八列(读者可自行验证)。■

要用卡诺图化简 n 变元的布尔函数，首先应画出合适大小的卡诺图。我们在积之和扩展中的小项对应的所有方格中放入 1，然后确定出函数的所有素隐含数。要做到这一点，我们寻找由 2^k 个聚簇方格(全包含 1)组成的块，其中 $1\leqslant k\leqslant n$。这些块对应于 $n-k$ 个文字的积。(练习 33 会与此相关。)而且，2^k 个方格(全包含 1)的块不包含在表示素隐含数的 2^{k+1} 个方格(全含 1)之中。

例 7 在化简五变元布尔函数的卡诺图中，有一表示两个文字之积的 8 个方格全是 1 的块，若它不包含在一个 16 方格全是 1 且表示单个文字的块中，则此块是素隐含的。■

一旦所有的素隐含数确定后，我们的目标是找出这些素隐含数的具如下性质的最小可能子集：子集中的素隐含数覆盖了卡诺图中所有包含 1 的单元。首先应选择本原素隐含数，然后增加其他素隐含数以确保覆盖图中所有为 1 的方格。当变元的数量较大时，这最后一步会极为复杂。

10.4.3 无需在意的条件

在某些电路中，由于输入值的一些组合从未出现过，所以我们关心电路对输入值的其他组合的输出，这使得我们在生产具有所需输出的电路时有很大自由，因为对于所有不出现的输入值的组合，其输出值可以任意选择。函数对于这种组合的值称为*无需在意条件*。在卡诺图中，对于那些其函数值可以任意选择的变元值组合，用 d 对其作记号。在化简过程中，如果输入值的组合在卡诺图中导致最大的块，则我们可以将其赋值 1。下面例子说明了这一点。

例 8 用二进制数字对十进制展开式进行编码的一种方法是：对此十进制展开式中的每一位，在编码的二进制展开式中用四个二进制数字对其编码，例如，873 的编码为 100001110011。十进制展开式的这种编码方式称为*十进制数二元编码*。因为有 16 个四位二进制数，但只有 10 个十进制数字，所以还有 6 个四位二进制数没有被用来对数位进行编码。假设现在需要构造一个电路使得：如果数位大于或等于 5，则输出 1；若数位小于 5，则输出 0。怎么仅用与门、或门和反相器来构造这个电路？

解 以 $F(w, x, y, z)$ 记此电路的输出，其中 $wxyz$ 是一个十进制数位的二进制表达式。F 的值如表 10-13 所示，图 10-22a 是 F 的卡诺图，其中的无须在意位置都是 d。我们可以将 d 包括在块中或者剔除出去，这样块就有很多可能的选择。例如，如果剔除所有的 d 方格，则形成块如图 10-22b 所示，所产生的表达式为 $w\bar{x}\bar{y}+\bar{w}xy+\bar{w}xz$；如果包括某些 d 而剔除其余的，则形成的块如图 10-22c 所示，且所产生的表达式为 $w\bar{x}+\bar{w}xy+x\bar{y}z$；最后，如果包括所有的块，且使用如图 10-22d 所示的块，则产生最简单的展开式，即 $F(x, y, z)=w+xy+xz$。■

表 10-13

数字	w	x	y	z	F
0	0	0	0	0	0
1	0	0	0	1	0
2	0	0	1	0	0
3	0	0	1	1	0
4	0	1	0	0	0
5	0	1	0	1	1
6	0	1	1	0	1
7	0	1	1	1	1
8	1	0	0	0	1
9	1	0	0	1	1

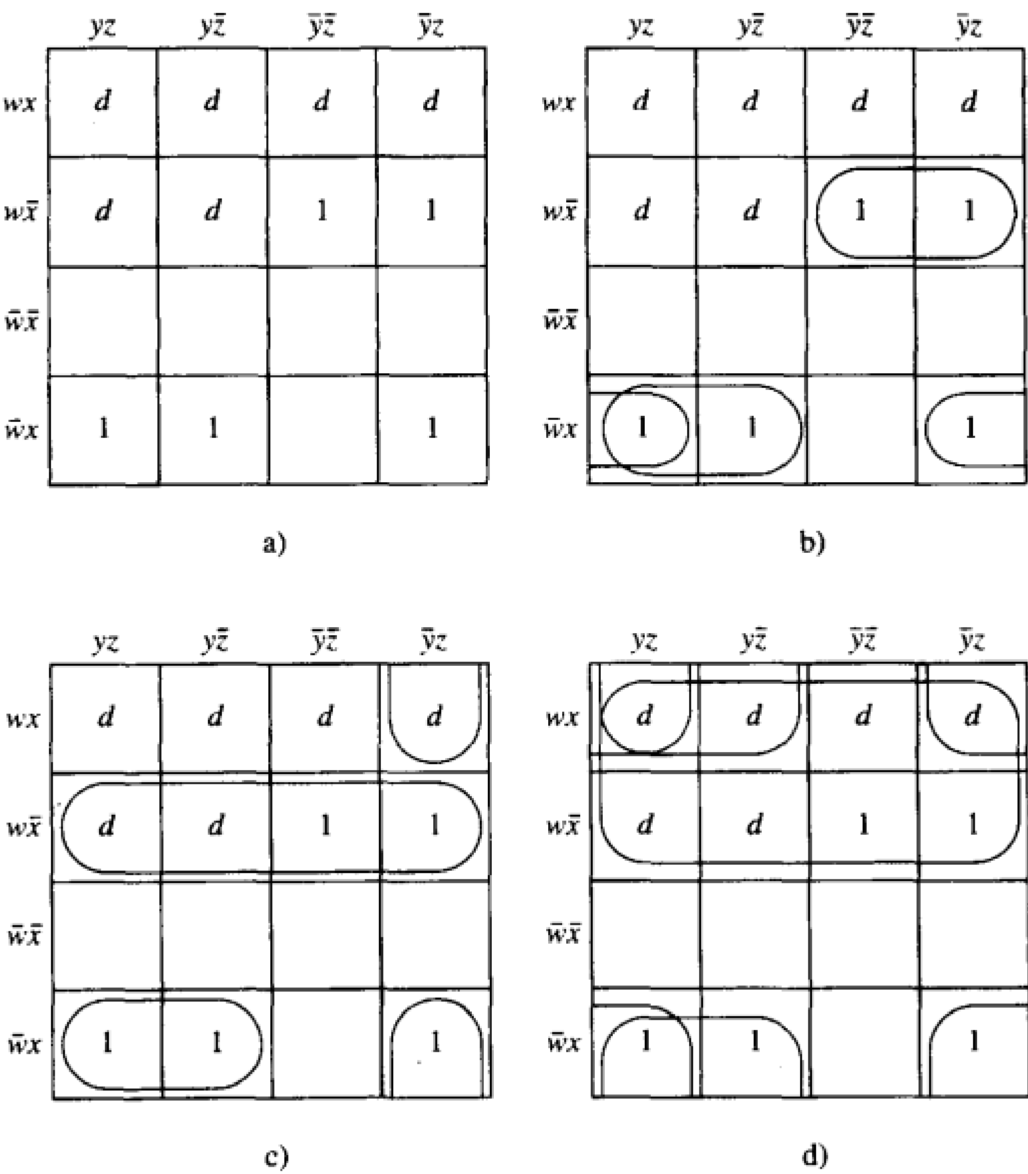


图 10-22 表明其无需在意条件位置的卡诺图

10.4.4 奎因-莫可拉斯基方法

我们已经看到，可以用卡诺图将布尔函数展开为形如布尔积之布尔和的极小表达式。但当变元超过四个时，卡诺图就变得难以使用，而且，卡诺图的使用还要依赖于用目测方法将项分成组。鉴于这些原因，需要可以机械化的过程来化简积之和展开式，奎因-莫可拉斯基方法就是这样一种过程，它可以用于含有任意多个变元的布尔函数。此法是由 W. V. 奎因和 E. J. 莫可拉斯基[⊖]于 20 世纪 50 年代提出的。它基本上由两部分组成，第一部分寻找可能包含在形如布尔积之布尔和的极小展开式中的候选项，第二部分才确定哪些项将真正使用。下面用例子来说明这个过程是怎样进行的。

例 9 下面说明怎么用奎因-莫可拉斯基方法寻找等价于 $xyz+x\bar{y}z+\bar{x}yz+\bar{x}\bar{y}z+\bar{x}\bar{y}\bar{z}$ 的极小展开式。

我们用一串二进制数位来表示此展开式中的小项。如果 x 出现，则第一位为 1；如果 $\bar{x}$ 出现，

⊖ **爱德华·莫可拉斯基(Edward J. McCluskey)**生于 1929 年，就读于 Bowdoin 学院和麻省理工学院，并于 1956 年在麻省理工学院获得电子工程学博士学位。他于 1955 年进入贝尔电话实验室，在那里一直待到 1959 年。从 1959 年到 1966 年，莫可拉斯基是普林斯顿大学的电子工程学教授，并在 1961～1966 年间在普林斯顿担任计算中心主任。他于 1967 年在斯坦福大学担任计算机科学和电子工程学的教授，并于 1969～1978 年间，担任数字系统实验室的主任。莫可拉斯基对计算机科学的许多领域有过研究，包括容错计算、计算机体系结构、测试和逻辑设计。他现在是斯坦福大学的可靠性计算中心的主任，还是 ACM(美国计算机协会)的会员。

则第一位为 0。如果 y 出现，则第二位为 1；如果 $\overline{y}$ 出现，则第二位为 0。如果 z 出现，则第三位为 1；如果 $\overline{z}$ 出现，则第三位为 0。然后根据对应数位串中 1 的个数来对这些项进行分组。这些信息如表 10-14 所示。

表 10-14

小项	数位串	1 的个数
xyz	111	3
$x\overline{y}z$	101	2
$\overline{x}yz$	011	2
$\overline{x}\overline{y}z$	001	1
$\overline{x}\overline{y}\overline{z}$	000	0

可以合并的小项只在一个文字处不同，所以，对于两个可以合并的小项，在表示它们的数位串中，1 的个数仅相差 1。当两个小项被合并成一个积时，这个积只含有两个文字。两个文字的积可以如下表示：以短划线来记不出现的变元。例如，数位串 101 和 001 所表示的小项 $x\overline{y}z$ 和 $\overline{x}\overline{y}z$ 可以合并成 $\overline{y}z$，而 $\overline{y}z$ 可以用数位串—01 表示。表 10-15 列出了所有可以合并的成对小项以及它们所产生的积。

下一步，对于由两个文字构成的积，如果两个这样的积能够合并，则将之合并成一个文字。两个这样的积能够合并的条件是：它们所包含的文字是两个相同变元的文字，并且只有其中一个变元的文字不一致。就表示这些积的串来说，它们必定在相同位置有一个短划线，且在其余的两个位置中必定有一个位置的内容不相同。我们可以将串—11 和—01 所表示的积 yz 和 $\overline{y}z$ 合并成 z，而 z 是由串——1 表示的。所有能够以这种方式合并的项如表 10-15 所示。

表 10-15

	项	数位串	步骤 1 项	步骤 1 数位串		步骤 2 项	步骤 2 数位串	
1	xyz	111	(1，2)	xz	1—1	(1，2，3，4)	z	——1
2	$x\overline{y}z$	101	(1，3)	yz	—11			
3	$\overline{x}yz$	011	(2，4)	$\overline{y}z$	—01			
4	$\overline{x}\overline{y}z$	001	(3，4)	$\overline{x}z$	0—1			
5	$\overline{x}\overline{y}\overline{z}$	000	(4，5)	$\overline{x}\overline{y}$	00—			

在表 10-15 中，我们还指出了哪些项可以用来形成更少文字的积，这些项不一定在极小展开式中。下一步是找出积的一个极小集合，使之可以用来表示此布尔函数。我们从那些还没有被用来形成更少文字之积的积着手。再下一步，我们构造表 10-16，通过合并原来项所形成的每一个候选积构成此表的行，原来的项构成列。如果积之和展开式中原来的项被用来形成这个候选积，则在相应的位置打上 X，此时称此侯选项**覆盖**了原来的小项。我们需要至少一个积，它覆盖原来的每一个小项。因而，一旦此表的某一列只有一个 X，则此 X 所在的行所对应的积必定被使用。从表 10-16 可以看出，z 和 $\overline{x}\overline{y}$ 都是必需的。所以，最后的答案是 $z+\overline{x}\overline{y}$。 ■

表 10-16

	xyz	$x\overline{y}z$	$\overline{x}yz$	$\overline{x}\overline{y}z$	$\overline{x}\overline{y}\overline{z}$
z	X	X	X	X	
$\overline{x}\overline{y}$				X	X

就像例 9 所说明的那样，奎因—莫可拉斯基方法用下面一系列步骤来化简一系列积之和展开式。

1)将由 n 个变元构成的每一个小项表示成一个长度为 n 的二进制数串，如果 x_i 出现，则此串的第 i 个位置为 1；如果 $\overline{x_i}$ 出现，则此串的第 i 个位置为 0。

2)根据串中 1 的个数将串分组。

3)确定所有这样 $n-1$ 个变元的积，它们可以取为此展开式中小项的布尔和。将能够合并的小项表示成二进位串，且这些串只在一个位置不相同。将这些 $n-1$ 个变元的积用如下的串表示：如果 x_i 出现在此积中，则此串的第 i 个位置为 1；如果$\overline{x_i}$出现，则此位置为 0；如果此积中没有涉及 x_i 的文字，则此位置为短划线。

4)确定所有这样 $n-2$ 个变元的积，它们可以取为在前一个步骤形成的 $n-1$ 个变元的积的布尔和。将能够合并的 $n-1$ 个变元的积表示成如下的串：在同一位置有一个短划线，且只在一个位置不相同。

5)只要可能，继续将布尔积合并成更少变元的积。

6)找到所有这样的布尔积：它们虽然出现，但还没有被用来形成少一个文字的布尔积。

7)找到这些布尔积的最小集合，使得这些积的和表示此布尔函数。这可以用如下方法来完成：构造一个表，列出哪些积覆盖了哪些小项。每一个小项必定被至少一个积覆盖。(这是此过程中的最困难部分，可用回溯法将其机械化。)

下面最后的例子说明了怎么用这个过程来化简四个变元的积之和展开式。

例 10 用奎因-莫可拉斯基法化简积之和展开式 $wxy\bar{z}+w\bar{x}yz+w\bar{x}y\bar{z}+\bar{w}xyz+\bar{w}x\bar{y}z+\bar{w}\bar{x}yz+\bar{w}\bar{x}\bar{y}z$

解 首先将小项表示成二进位串，然后根据串中 1 的个数来对项进行分组，如表 10-17 所示。表 10-18 给出了所有的布尔积，它们可以取为这些积的布尔和。

表 10-17

项	数位串	1 的个数
$wxy\bar{z}$	1110	3
$w\bar{x}yz$	1011	3
$\bar{w}xyz$	0111	3
$w\bar{x}y\bar{z}$	1010	2
$\bar{w}x\bar{y}z$	0101	2
$\bar{w}\bar{x}yz$	0011	2
$\bar{w}\bar{x}\bar{y}z$	0001	1

表 10-18

			步骤 1			步骤 2		
	项	数位串		项	数位串		项	数位串
1	$wxy\bar{z}$	1110	(1, 4)	$wy\bar{z}$	1−10	(3, 5, 6, 7)	$\bar{w}z$	0−−1
2	$w\bar{x}yz$	1011	(2, 4)	$w\bar{x}y$	101−			
3	$\bar{w}xyz$	0111	(2, 6)	$\bar{x}yz$	−011			
4	$w\bar{x}y\bar{z}$	1010	(3, 5)	$\bar{w}xz$	01−1			
5	$\bar{w}x\bar{y}z$	0101	(3, 6)	$\bar{w}yz$	0−11			
6	$\bar{w}\bar{x}yz$	0011	(5, 7)	$\bar{w}\bar{y}z$	0−01			
7	$\bar{w}\bar{x}\bar{y}z$	0001	(6, 7)	$\bar{w}\bar{x}z$	00−1			

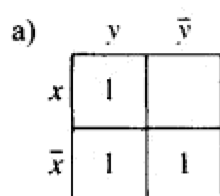
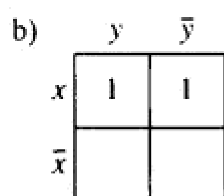
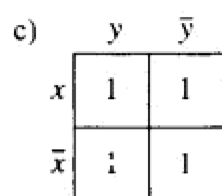
没有被用来形成更少变元之积的只有$\bar{w}z$、$wy\bar{z}$、$w\bar{x}y$ 和$\bar{x}yz$。表 10-19 表明了每个这样的积覆盖的小项。为覆盖这些小项，必须包括$\bar{w}z$ 和 $wy\bar{z}$，因为它们是分别覆盖$\bar{w}xyz$ 和 $wxy\bar{z}$ 的唯一的积。一旦这两个积包括进来，我们就可以看到，只有其中的一个是必要的。因而，$\bar{w}z+wy\bar{z}+w\bar{x}y$ 或者$\bar{w}z+wy\bar{z}+\bar{x}yz$ 都可以看做是最后答案。■

表　10-19

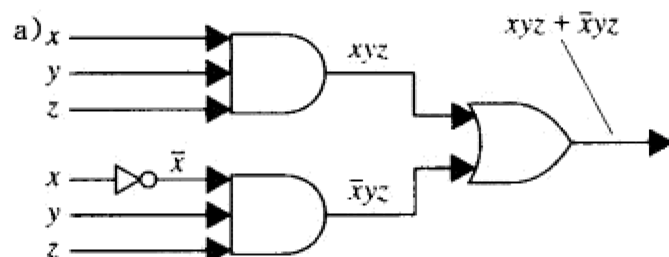
	$wxy\bar{z}$	$w\bar{x}yz$	$\bar{w}xyz$	$w\bar{x}y\bar{z}$	$\bar{w}x\bar{y}z$	$\bar{w}\bar{x}yz$	$\bar{w}\bar{x}\bar{y}z$
$\bar{w}z$			X		X	X	X
$wy\bar{z}$	X			X			
$w\bar{x}y$		X		X			
$\bar{x}yz$		X				X	

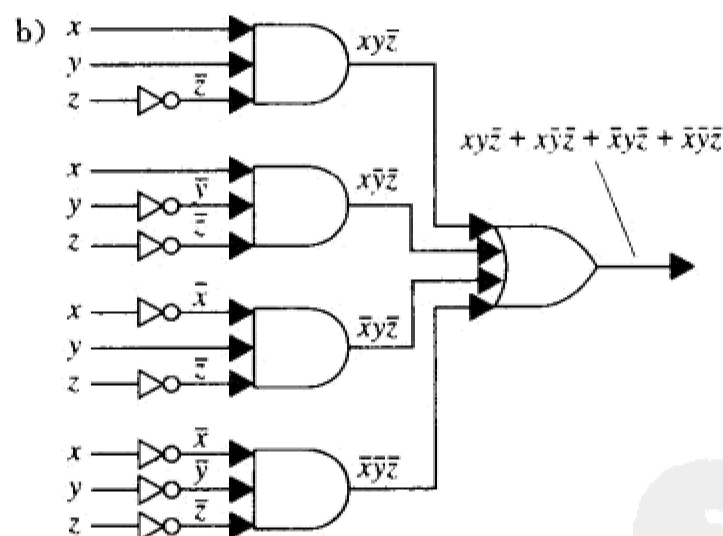
练习

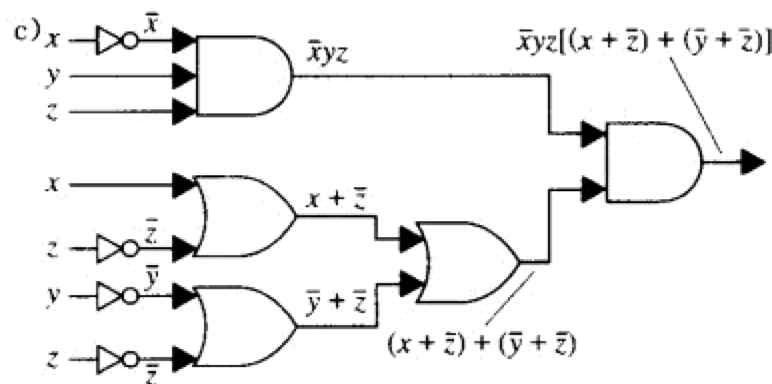
1. a)画出二变元函数的卡诺图，并在表示$\bar{x}y$的方格中放置1。
   b)与上述方格相邻的方格所表示的小项是什么？
2. 寻找下列每个卡诺图所表示的积之和展开式。

a) 
b) 
c) 

3. 画出下列两个变元的积之和展开式的卡诺图：
   a)$x\bar{y}$　　b)$xy+\bar{x}\bar{y}$　　c)$xy+x\bar{y}+\bar{x}y+\bar{x}\bar{y}$
4. 用卡诺图找出下列关于变元x和y的布尔函数的极小展开式，且此展开式具有布尔积之布尔和的形式。
   a)$\bar{x}y+\bar{x}\bar{y}$　　b)$xy+x\bar{y}$　　c)$xy+x\bar{y}+\bar{x}y+\bar{x}\bar{y}$
5. a)画出三变元函数的卡诺图，并在表示$\bar{x}y\bar{z}$的方格里放置1。
   b)与上述方格相邻的方格的极小项是什么？
6. 对于下列电路图，用卡诺图画出具有相同输出的更简单的电路图。

a) 

b) 

c) 

7. 画出下列三变元积之和展开式的卡诺图：
   a)$x\bar{y}\bar{z}$　　b)$\bar{x}yz+\bar{x}\bar{y}\bar{z}$　　c)$xyz+xy\bar{z}+\bar{x}y\bar{z}+\bar{x}\bar{y}z$
8. 构造$F(x, y, z)=xz+yz+xy\bar{z}$的卡诺图。使用此卡诺图找出$F(x, y, z)$的隐含数、素隐含数和本原素隐含数。

9. 构造 $F(x, y, z)=x\overline{z}+xyz+y\overline{z}$ 的卡诺图。使用此卡诺图找出 $F(x, y, z)$ 的隐含数、素隐含数和本原素隐含数。

10. 画一个 3 立方体 Q_3，用布尔变元 x、y 和 z 组成的极小项标记每一个顶点，这些项与顶点表示的位串关联。对这些变元中的每一个文字，指出表示这个文字且是 Q_3 的子图的 2 立方体 Q_2。

11. 画一个 4 立方体 Q_4，用布尔变元 w、x、y 和 z 组成的极小项标记每一个顶点，这些项与顶点表示的位串关联。对这些变元中的每一个文字，指出哪个 3 立方体 Q_3 表示这个文字且是 Q_4 的子图。指出哪个 2 立方体 Q_2 表示积 wz、$\overline{x}y$ 和 $\overline{y}\,\overline{z}$ 且是 Q_4 的子图。

12. 用卡诺图找出下列关于变元 x、y 和 z 的函数的一个极小展开式，且此展开式具有布尔积之布尔和的形式。

    a) $\overline{x}yz+\overline{x}\,\overline{y}z$

    b) $xyz+xy\overline{z}+\overline{x}yz+\overline{x}y\overline{z}$

    c) $xy\overline{z}+x\overline{y}z+x\overline{y}\,\overline{z}+\overline{x}yz+\overline{x}\,\overline{y}z$

    d) $xyz+x\overline{y}z+x\overline{y}\,\overline{z}+\overline{x}yz+\overline{x}y\overline{z}+\overline{x}\,\overline{y}\,\overline{z}$

13. a) 画出四变元函数的卡诺图，并在 $\overline{w}xy\overline{z}$ 所表示的方格里填入 1。

    b) 与上述方格相邻的方格的极小项是什么？

14. 用卡诺图找出下列关于变元 w、x、y 和 z 的函数的一个极小展开式，且此展开式具有布尔积之布尔和的形式。

    a) $wxyz+wx\overline{y}z+wx\overline{y}\,\overline{z}+w\overline{x}y\overline{z}+w\overline{x}\,\overline{y}z$

    b) $wxy\overline{z}+wx\overline{y}z+w\overline{x}yz+\overline{w}x\overline{y}z+\overline{w}\,\overline{x}y\overline{z}+\overline{w}\,\overline{x}\,\overline{y}z$

    c) $wxyz+wxy\overline{z}+wx\overline{y}z+w\overline{x}\,\overline{y}z+w\overline{x}\,\overline{y}\,\overline{z}+\overline{w}x\overline{y}z+\overline{w}\,\overline{x}y\overline{z}+\overline{w}\,\overline{x}\,\overline{y}z$

    d) $wxyz+wxy\overline{z}+wx\overline{y}z+w\overline{x}yz+w\overline{x}y\overline{z}+\overline{w}xyz+\overline{w}\,\overline{x}yz+\overline{w}\,\overline{x}y\overline{z}+\overline{w}\,\overline{x}\,\overline{y}z$

15. 在表示五变元布尔函数的卡诺图中，找出对应于下列积的方格。

    a) $x_1x_2x_3x_4$　　b) $\overline{x}_1x_3x_5$　　c) x_2x_4

    d) $\overline{x}_3\overline{x}_4$　　e) x_3　　f) $\overline{x}_5$

16. 在六变元布尔函数的卡诺图中，表示 x_1，$\overline{x}_1x_6$，$\overline{x}_1x_2\overline{x}_6$，$x_2x_3x_4x_5$ 和 $x_1\overline{x}_2x_4\overline{x}_5$ 分别需要多少方格？

17. a) 六变元函数的卡诺图具有多少个方格？

    b) 在六变元函数的卡诺图中，对于任意给定的一个方格，有多少个方格与之相邻？

18. 证明：在五变元布尔函数的卡诺图中，表示极小项的方格恰在一个文字上不同当且仅当这些极小项相邻，或者顶行和底行的方格相邻，第一列和第八列的方格相邻，第一和第四列，第二和第七列，第三和第六列，第五和第八列的方格相邻。

19. 在六变元布尔函数的 4×16 卡诺图中，若用格雷码 1111，1110，1010，1011，1001，1000，0000，0001，0011，0010，0110，0111，0101，0100，1100，1101 标记列，用 11，10，00，01 标记行，则哪些行和列应当相邻才可使得：恰在一个文字上不同的表示极小项的方格相邻？

*20. 用卡诺图找出下列函数的极小展开式，使得此展开式具有布尔积之布尔和的形式，这些函数满足：其输入为十进制数字的二进制编码，其输出为 1 当且仅当对应于输入的数为

    a) 奇数　　b) 不可由 3 整除　　c) 不是 4、5 或 6

*21. 假设一个委员会中有五个成员，其中的施密斯和琼斯的投票总与马库斯的投票相反。试用这个投票关系设计一个电路，实现此委员会的多数表决。

22. 使用奎因-莫可拉斯基法化简例 3 中的积之和展开式。

23. 使用奎因-莫可拉斯基法化简练习 12 中的积之和展开式。

24. 使用奎因-莫可拉斯基法化简例 4 中的积之和展开式。

25. 使用奎因-莫可拉斯基法化简练习 14 中的积之和展开式。

*26. 试解释怎么用卡诺图方法简化三个变元的和之积展开式。(提示：用 0 来标记展开式的极大项，然后构造极大项的块。)

27. 用练习 26 的方法化简和之积展开式 $(x+y+z)(x+y+\overline{z})(x+\overline{y}+\overline{z})(x+\overline{y}+z)(\overline{x}+y+z)$。

*28. 在圆环面上画出四个变元的 16 个极小项的卡诺图。

29. 用或门、与门和反相器构造一个电路，使得：当输入的十进制数字可以被 3 整除时输出 1，否则输出 0。其中：输入的十进制数字是二进制编码的十进制展开式。

对于练习 30～32，在所给的卡诺图中，d 表示无须在意条件。试找出它们的极小积之和展开式。

30.

	yz	$y\overline{z}$	$\overline{y}\overline{z}$	$\overline{y}z$
wx	d	1	d	1
$w\overline{x}$		d	d	
$\overline{w}\overline{x}$		d	1	
$\overline{w}x$		1	d	

31.

	yz	$y\overline{z}$	$\overline{y}\overline{z}$	$\overline{y}z$
wx	1			1
$w\overline{x}$		d	1	
$\overline{w}\overline{x}$		1	d	
$\overline{w}x$	d			d

32.

	yz	$y\overline{z}$	$\overline{y}\overline{z}$	$\overline{y}z$
wx		d	d	1
$w\overline{x}$	d	d	1	d
$\overline{w}\overline{x}$				
$\overline{w}x$	1	1	1	d

33. 证明：k 个文字的积对应于 n 立方体 Q_n 的 2^{n-k} 维子立方体，其中立方体的顶点对应于标识顶点的位串表示的小项，如 8.2 节例 7 的描述。

关键术语和结果

术语

布尔变元：只取 0 或 1 值的变元。

$\overline{x}$(x 的余)：一个表达式，当 x 取值 0 时，它取值 1；当 x 取值 1 时，它取值 0。

$x \cdot y$(或 xy)(x 与 y 的布尔积或合取)：一个表达式，当 x 和 y 都取值 1 时，它取值 1；否则取值 0。

$x+y$(x 与 y 的布尔和或析取)：一个表达式，当 x 或 y 取值 1 时，或者当 x 和 y 都取值 1 时，它取值 1；否则取值 0。

布尔表达式：如下递归得到的表达式：0，1，x_1，…，x_n 是布尔表达式；且如果 E_1 和 E_2 是布尔表达式，则 $\overline{E_1}$、(E_1+E_2) 和 (E_1E_2) 也是布尔表达式。

布尔表达式的对偶：通过交换＋号和·号、0 和 1 得到的表达式。

n 度布尔函数：从 B^n 到 B 的函数，其中 $B=\{0, 1\}$。

布尔代数：具有两个二元运算 ∨ 和 ∧、元素 0 和 1、一元补运算 $^-$ 的集合，它满足同一律、支配律、结合律、交换律和分配律。

布尔变元的文字 x：或者为 x，或者为 $\overline{x}$。

x_1，x_2，…，x_n 的小项：布尔积 $y_1y_2\cdots y_n$，其中每个 y_i 或为 x_i 或为 $\overline{x_i}$。

积之和展开式(或析取范式)：形如小项之析取的布尔函数的表示。

函数完全：布尔运算的一个集合称为是函数完全的，如果每个布尔函数都能由这些布尔运算表示。

$x \mid y$(或 x NAND y)：一个表达式，当 x 和 y 都取值 1 时，它取值 0；否则取值 1。

$x \downarrow y$(或 x NOR y)：一个表达式，当 x 或 y 取值 1 时，或 x 和 y 都取值 1 时，它取值 0；否则取值 1。

反相器：一种装置，它以布尔变元的值作为输入，产生输入的补。

或门：一种装置，它以两个或更多布尔变元的值作为输入，输出它们的布尔和。

与门：一种装置，它以两个或更多布尔变元的值作为输入，输出它们的布尔积。

半加法器：一种电路，它将两个二进制数字相加，产生一个和位与一个进位。

全加法器：一种电路，它将两个二进制数字及一个进位相加，产生一个和位与一个进位。

n 个变元的卡诺图：被分成 2^n 多个方格的矩形，每个方格表示这些变元的一个小项。

布尔函数的最小化：将布尔函数表示为最小的文字积，这些积包含表示这个布尔函数的积的可能的所有和中最少的文字。

布尔函数的隐含：如果文字积为 1，那么布尔函数的值为 1。具有这个性质的文字积。

布尔函数的素隐含：布尔函数的隐含文字积和不删除文字得到的积也是这个函数的隐含。

布尔函数的基本素隐含：布尔函数的素隐含必须包括在这个函数的最小化中。

结果

布尔代数中的恒等式(见 10.1 节的表 10-5)。

对于布尔表达式表示的布尔函数间的任意等式，如将等式的两边取对偶，则等式依然成立。

每个布尔函数都可由积之和展开式表示。

集合{+,⁻}和{·,⁻}都是函数完全的。

集合{↓}和{|}都是函数完全的。

使用卡诺图来极小化布尔表达式。

使用奎因-莫可拉斯基法来极小化布尔表达式。

复习题

1. 给出 n 度的布尔函数的定义。
2. 有多少个 2 度布尔函数？
3. 给出布尔表达式集合的递归定义。
4. a)什么是布尔表达式的对偶？
   b)什么是对偶原理？怎么应用它找到关于布尔表达式的新等式？
5. 试解释怎么构造一个布尔函数的积之和展开式。
6. a)“由运算符构成的一个集合是函数完全的”是什么含义？
   b)集合{+，·}是函数完全的吗？
   c)有没有单运算符构成的集合是函数完全的？
7. 试解释怎么用或门、与门和反相器构造一个电路，它用两个开关控制一盏灯。
8. 用或门、与门和反相器构造一个半加法器。
9. 是否有这样一种逻辑门：用之可以构造或门、与门和反相器所能构造的所有电路？
10. a)解释怎么用卡诺图来化简三个布尔变元的积之和展开式。
    b)用卡诺图化简积之和展开式 $xyz+x\bar{y}z+x\bar{y}\bar{z}+\bar{x}yz+\bar{x}\bar{y}\bar{z}$。
11. a)解释怎么用卡诺图来化简四个布尔变元的积之和展开式。
    b)用卡诺图化简积之和展开式
    $$wxyz+wxy\bar{z}+wx\bar{y}z+wx\bar{y}\bar{z}+w\bar{x}yz+w\bar{x}\bar{y}z+\bar{w}xyz+\bar{w}\bar{x}yz+\bar{w}\bar{x}y\bar{z}$$
12. a)什么是无须在意条件？
    b)试解释怎么用无须在意条件由或门、与门和反相器构造这样一个电路：当十进制数字大于等于 6 时输出 1，当这个数字小于 6 时输出 0。
13. a)试解释怎么用奎因-莫克拉斯基方法来化简积之和展开式。
    b)用这个方法化简 $xy\bar{z}+x\bar{y}\bar{z}+\bar{x}y\bar{z}+\bar{x}\bar{y}z+\bar{x}\bar{y}\bar{z}$。

补充练习

1. 对于布尔变元 x、y 和 z 的哪些值
   a) $x+y+z=xyz$　　b) $x(y+z)=x+yz$　　c) $\bar{x}\bar{y}\bar{z}=x+y+z$

2. 设 x 和 y 属于{0, 1}。如果存在{0, 1}中的值 z 使得下式之一成立，有无必要 $x=y$ 成立？

a) $xz=yz$　　b) $x+z=y+z$　　c) $x\oplus z=y\oplus z$

d) $x\downarrow z=y\downarrow z$　　e) $x\mid z=y\mid z$

布尔函数 F 称为是自对偶的当且仅当 $F(x_1,\cdots,x_n)=\overline{F(\overline{x_1},\cdots,\overline{x_n})}$。

3. 下列函数哪些是自对偶的？

a) $F(x,y)=x$　　b) $F(x,y)=xy+\overline{x}\,\overline{y}$　　c) $F(x,y)=x+y$　　d) $F(x,y)=xy+\overline{x}y$

4. 试给出一个三变元自对偶布尔函数的例子。

*5. 有多少个 n 度布尔函数是自对偶的？

在 n 度布尔函数构成的集合上，定义关系≤使得：$F\leqslant G$ 当且仅当若 $F(x_1,x_2,\cdots,x_n)=1$ 就有 $G(x_1,x_2,\cdots,x_n)=1$。

6. 对于下列函数对，确定是否有 $F\leqslant G$ 或 $G\leqslant F$。

a) $F(x,y)=x$，$G(x,y)=x+y$　　b) $F(x,y)=x+y$，$G(x,y)=xy$

c) $F(x,y)=\overline{x}$，$G(x,y)=x+y$

7. 设 F 和 G 是 n 度布尔函数，证明：

a) $F\leqslant F+G$　　b) $FG\leqslant F$。

8. 设 F、G 和 H 都是 n 度布尔函数。证明：$F+G\leqslant H$ 当且仅当 $F\leqslant H$ 且 $G\leqslant H$。

*9. 证明≤关系是 n 度布尔函数集合上的一个偏序关系。

*10. 画出由 16 个 2 度布尔函数(如 10.1 节表 10-3 所示)的集合构成的幂集在偏序≤下的哈斯图。

*11. 对于下列每个等式，或者证明其为恒等式，或者找到变元的一组值使之不成立。

a) $x\mid(y\mid z)=(x\mid y)\mid z$　　b) $x\downarrow(y\downarrow z)=(x\downarrow y)\downarrow(x\downarrow z)$

c) $x\downarrow(y\mid z)=(x\downarrow y)\mid(x\downarrow z)$

定义布尔运算符⊙如下：$1\odot 1=1$，$1\odot 0=0$，$0\odot 1=0$，$0\odot 0=1$。

12. 证明 $x\odot y=xy+\overline{x}\,\overline{y}$。

13. 证明 $x\odot y=\overline{(x\oplus y)}$。

14. 证明下列各等式成立：

a) $x\odot x=1$　　b) $x\odot\overline{x}=0$　　c) $x\odot y=y\odot x$。

15. $(x\odot y)\odot z=x\odot(y\odot z)$是否总成立？

*16. 确定集合{⊙}是不是函数完全的。

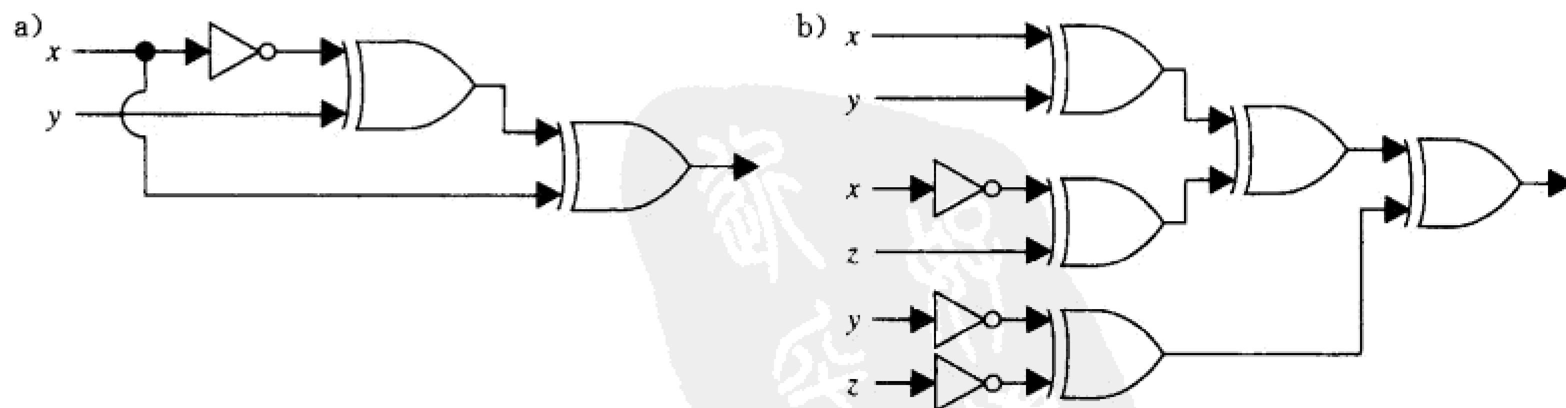
*17. 在 16 个两变元 x 和 y 的布尔函数中，有多少个能够用下列运算集、变元 x 和 y 以及值 0 和 1 表示？

a) {¯}　　b) {·}　　c) {+}　　d) {·, +}

异或门的记号如下，它从 x 和 y 产生输出 $x\oplus y$。

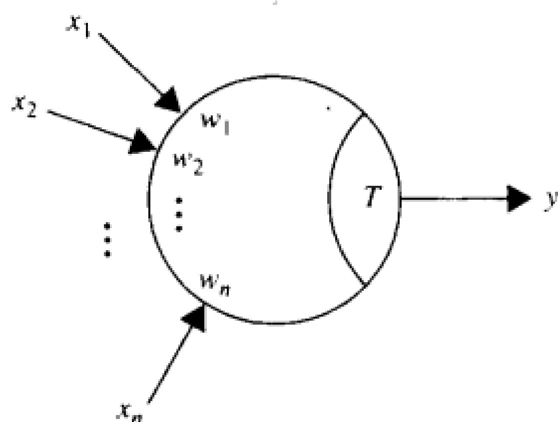


18. 确定下列电路 a) 和 b) 的输出。

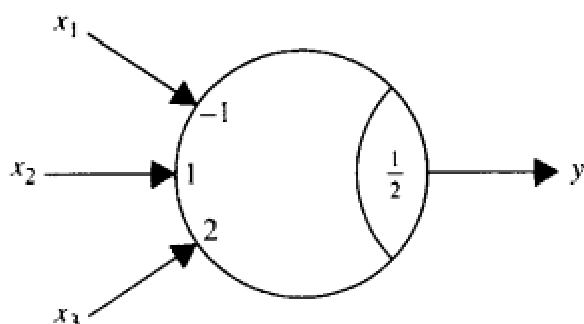


19. 如果除了或门、与门和反相器之外，还可以使用异或门，说明怎么用比 10.3 节图 10-9 中所用的更少的门来构造一个半加法器。
20. 试设计一个电路来确定：在一个四人委员会中，是否有三人或更多的人就某事投了赞成票，其中的每个人用一个开关来投票。

给定布尔变元 x_1，x_2，…，x_n 的一组输入值，阈值门产生输出 y，其中 y 为 0 或 1。每个阈值门都有一个阈值 T 以及一组权 w_1，w_2，…，w_n，其中 T 及 w_1，w_2，…，w_n 都是实数。阈值门的输出 y 是 1 当且仅当 $w_1x_1+w_2x_2+\cdots+w_nx_n\geq T$。具有阈值 T 及权 w_1，w_2，…，w_n 的阈值门如下图所示。阈值门对于神经生理学和人工智能的建模都非常有用。



21. 阈值门表示了一个布尔函数。试找出由下面阈值门表示的布尔函数的布尔表达式。



22. 能够由阈值门表示的布尔函数称为阈值函数。证明下列每个函数都是阈值函数。

a) $F(x)=\overline{x}$　　b) $F(x, y)=x+y$　　c) $F(x, y)=xy$

d) $F(x, y)=x \mid y$　　e) $F(x, y)=x\downarrow y$　　f) $F(x, y, z)=x+yz$

g) $F(w, x, y, z)=w+xy+z$　　h) $F(w, x, y, z)=wxz+x\overline{y}z$

*23. 证明 $F(x, y)=x\oplus y$ 不是阈值函数。

*24. 证明 $F(w, x, y, z)=wx+yz$ 不是阈值函数。

计算机题目

写出具有下列输入和输出的程序。

1. 给定两个布尔变元 x 和 y 的值，计算 $x+y$、xy、$x\oplus y$、$x\mid y$ 和 $x\downarrow y$ 的值。
2. 构造一个表，列出所有 256 个 3 度布尔函数的值。
3. 给定一个 n 元布尔函数的所有值，其中 n 是正整数，构造这个函数的积之和展开式。
4. 给定一个布尔函数值的列表，只用运算 · 和 ¯ 表示这个函数。
5. 给定一个布尔函数值的列表，只用运算 + 和 ¯ 表示这个函数。

*6. 给定一个布尔函数值的列表，只用运算 | 表示这个函数。

*7. 给定一个布尔函数值的列表，只用运算 ↓ 表示这个函数。

8. 给定一个 3 度布尔函数值的列表，构造它的卡诺图。
9. 给定一个 4 度布尔函数值的列表，构造它的卡诺图。

**10. 给定一个布尔函数值的列表，用奎因-莫可拉斯基方法寻找这个函数的极小积之和表示。

11. 对于一个阈值门和 n 个布尔变元的值，给定它的阈值和一组权，确定这个门的输出。

12. 给定一个正整数，构造一个 n 元随机布尔表达式，且为析取范式。

计算和研究

利用计算程序或你所编写的程序来做下列练习。

1. 计算 7、8、9、10 度布尔函数的个数。
2. 构造 3 度布尔函数的表。
3. 构造 4 度布尔函数的表。
4. 将每个不同的三元布尔表达式表示成仅含与非运算符的析取范式，所使用的与非运算符越少越好。所需与非运算符的最大数量是多少?
5. 将每个不同的布尔表达式表示成含有四个变元和仅含或非运算符的析取范式，所使用的或非运算符越少越好。所需或非运算符的最大数量是多少?
6. 随机生成 10 个不同的四变元布尔表达式，判断使用奎因-莫可拉斯基方法简化它们所需的平均步骤数。
7. 随机生成 10 个不同的五变元布尔表达式，判断使用奎因-莫可拉斯基方法简化它们所需的平均步骤数。

写作题目

用课本以外的资料解决下列问题，并写成短文。

1. 描述一些早期设计的、用来解逻辑问题的机器，如印刷示范器(Stanhope Demonstrator)、杰文(Jevons)的逻辑机以及马昆德机器(Marquand Machine)。
2. 解释组合电路与顺序电路间的差别，然后解释怎么用触发器构造顺序电路。
3. 定义移位寄存器，且讨论怎么使用移位寄存器。说明怎么用触发器和逻辑门构造移位寄存器。
4. 说明怎么用逻辑门构造乘法器。
5. 找出逻辑门的物理构造。讨论在构造电路时，是否要用到与非门和或非门。
6. 解释怎么用相关性记号描述复杂的开关电路。
7. 描述怎么用乘法器构造开关电路。
8. 以用阈值门构造半加法器和全加法器为例，解释用阈值门构造开关电路的优点。
9. 描述无危险开关电路的概念，并给出一些设计这种电路的原则。
10. 解释怎么用卡诺图将五元或六元函数极小化。
11. 讨论用于极小化布尔函数的新方法(如 Espresso 方法)的思想。解释怎么用这些方法来解决高达 25 元的布尔函数最小化问题。
12. 描述 n 元布尔函数的函数分解的含义。讨论将布尔函数分解为元数更少的布尔函数的过程。



第11章 计算模型

计算机能够执行许多任务。提出一个任务后就存在两个问题：第一，它能否由计算机来完成？一旦知道这第一个问题的答案是肯定的，就会问第二个问题，怎么执行这个任务？计算模型就是用来帮助回答这两个问题的。

下面将讨论三种类型的计算模型：文法、有限状态机和图灵机。文法是用来产生一种语言中的词，并且确定一个词是否属于一个语言。文法产生的形式语言不仅可以作为自然语言的模型，如英语，还可以作为程序设计语言的模型，如 Pascal、Fortran、Prolog、C 和 Java。特别是，文法在编译器的理论和构造中极为重要。在 20 世纪 50 年代，美国语言学家诺姆·乔姆斯基(Noam Chomsky)首先使用了下面将要讨论的文法。

在建模中，使用着各种类型的有限状态机。每个有限状态机都有一个状态集合(包含初始状态)和一个输入字母表，还有一个转移函数，它对每个由状态和输入构成的对，指定下一个状态。有限状态机的状态使得它具有有限的存储能力。有些有限状态机对每个转移产生一个输出符号。这类机器可以用作许多种机器的模型，如自动售货机、延迟机、二进制数加法器和语言识别器。我们还将讨论没有输出但具有终结状态的有限状态机，这样的机器广泛用于语言的识别，它们所识别的串是这样的串；由初始状态运行到终结状态。文法和有限状态机的概念具有紧密的联系，我们将要刻画有限状态机所能识别的集合的特征，并且证明这些集合恰恰就是某种类型文法所产生的集合。

然后引入图灵机的概念。我们将说明怎么用图灵机来识别集合，还要说明怎么用图灵机来计算数论函数。最后讨论图灵–丘奇论题：每个能行的计算都可由图灵机来完成。

11.1 语言和文法

11.1.1 引言

英语中，单词能以各种方式进行组合，单词的哪些组合可以构成有效句子是由英语文法确定的。例如：the frog writes neatly(青蛙的字写得很匀称)是一个有效的句子，因为它是由一个名词短语 the frog 接一个动词短语 writes neatly 构成的，其中名词短语 the frog 是由冠词 the 和名词 frog 组成的，动词短语 writes neatly 是由动词 writes 和副词 neatly 组成的。我们并不在意这是一个毫无意义的句子，因为我们只关心句子的*语法*，或者说形式，而不在意它的*语义*，或者说含义。我们也要指出，词的组合 swims quickly mathematics 不是有效的句子，因为它不符合英语文法规则。

自然语言(即口头语言)，如英语、法语、德语或西班牙语，都极为复杂。事实上，对一个自然语言，看起来不大可能说出它的所有语法规则。将一个语言自动翻译成另一个语言的研究引出了*形式语言*的概念。与自然语言不同，形式语言是由一组意义明确的语法规则定义的，语法规则不仅对于语言学和自然语言的研究十分重要，而且对于程序设计语言的研究也很重要。

我们要用文法描述形式语言的句子。在程序设计语言的应用中，经常出现两类问题：(1)怎么能够确定一组单词是否组合成了形式语言的一个有效句子？(2)怎么才能产生形式语言的一个有效句子。在考虑这两类问题时，文法的使用十分有益。

在给出文法的技术定义之前，先描述文法的一个例子，这个例子产生英语的一个子集，此英

语子集是用下列规则定义的，这些规则描述怎么产生有效的句子。这些规则是：

1. 句子是由一个名词短语后接一个动词短语形成的；
2. 名词短语由一个冠词接一个形容词再接一个名词组成，或者
3. 名词短语由一个冠词接一个名词组成；
4. 动词短语由一个动词接一个副词组成，或者
5. 动词短语由一个动词组成；
6. 冠词是 *a*，或者
7. 冠词是 *the*；
8. 形容词是 *large*，或者
9. 形容词是 *hungry*；
10. 名词是 *rabbit*，或者
11. 名词是 *mathematician*；
12. 动词是 *eats*，或者
13. 动词是 *hops*；
14. 副词是 *quickly*，或者
15. 副词是 *wildly*。

从这些规则出发，使用一系列替代直到不能再应用规则，就能形成一个有效的句子。例如，沿着下列替代序列就能得到一个有效句子：

句子
名词短语　动词短语
冠词　形容词　名词　动词短语
冠词　形容词　名词　动词　副词
the 形容词　名词　动词　副词
the large 名词　动词　副词
the large rabbit 动词　副词
the large rabbit hops 副词
the large rabbit hops quickly

容易看出，下面都是有效句子：*a hungry mathematician eats wildly*，*a large mathematician hops*，*the rabbit eats quickly*，等等。也可以看出，*the quickly eats mathematican* 不是有效句子。

11.1.2 短语结构文法

在给出文法的形式定义之前，先引入一个小术语。

定义 1 词汇表(字母表)V 是元素的一个有限的非空集合，其元素称为符号。V 上一个词(或句子)是由 V 中元素组成的有限长度的串。空串(或零串)是没有符号的串，记为 λ。V 上所有词的集合记为 V^*。V 上的一个语言是 V^* 的一个子集。

注意 空串 λ 是不包含任何符号的串。它不同于空集 $\varnothing$。因而 $\{\lambda\}$ 是仅包含一个串的集合，此串为空串。

可以用多种方式来指明语言。一种方式是列出语言中的所有词；还有一种方式是给出一些标准，使得一个词要在这个语言中，就必须满足这些标准。本节将描述另一种重要方式：使用文法，如使用本节引言中给出的规则集合。为产生词，文法提供一个由各种类型符号组成的集合和

一个由规则组成的集合。更确切地说，文法有一个*词汇表* V，V 是一个由符号组成的集合，语言中的成分就是这些符号导出的。词汇表中的某些元素不能由其他符号替换，这些元素称为*终结符*；词汇表中的其他成分可以用其他符号替换，它们被称为*非终结符*。终结符和非终结符集合通常分别记为 T 和 N。在本节引言所给的例子中，终结符集是{*a*，*the*，*rabbit*，*mathematician*，*hops*，*eats*，*quickly*，*wildly*}，非终结符集是{*句子*，*名词短语*，*动词短语*，*形容词*，*冠词*，*名词*，*动词*，*副词*}。词汇表中有一个特殊的元素，我们总是从这个特殊元素开始定义其他符号，此符号称为*初始符*，记为 S。在引言的例子中，初始符是*句子*。由词汇表 V 中元素构成的所有串的集合记为 V^*。指明 V^* 中的串能被什么样的串替代的规则称为文法的*产生式*，指明 w_0 可以替换为 w_1 的产生式记为 $w_0 \to w_1$。在本节引言所给的文法中，我们列举了所有产生式。如果使用刚才定义的记号，其中第一个产生式为*句子*→*名词短语　动词短语*。我们以下列定义作为小结。

定义 2　一个*短语结构文法*　$G=(V, T, S, P)$由下列四部分组成：词汇表 V，由 V 的所有终结符组成的 V 的子集合 T，V 的初始符 S，以及产生式集合 P。集合 $V-T$ 记为 N，N 中的元素称为*非终结符*。P 中的每个产生式的左边必须至少包含一个非终结符。

例 1　设 $G=(V, T, S, P)$，其中：$V=\{a, b, A, B, S\}$，$T=\{a, b\}$，S 是初始符，$P=\{S\to ABa, A\to BB, B\to ab, AB\to b\}$。则 G 是一个短语结构文法的例子。■

我们对短语结构文法的产生式所产生的词感兴趣。

定义 3　设 $G=(V, T, S, P)$是一个短语结构文法，$w_0 = lz_0r$（即 l，z_0 和 r 的连接）和 $w_1 = lz_1r$ 是 V 上的串。若 $z_0 \to z_1$ 是 G 的一个产生式，则称由 w_0 *可直接派生* w_1，记为 $w_0 \Rightarrow w_1$。如果 V 上的串 w_0，w_1，…，w_n（$n \geq 0$）满足 $w_0 \Rightarrow w_1$，$w_1 \Rightarrow w_2$，…，$w_{n-1} \Rightarrow w_n$，则称由 w_0 *可派生* w_n，记为 $w_0 \stackrel{*}{\Rightarrow} w_n$。由 w_0 得到 w_n 的序列称为*派生*。

例 2　在例 1 的文法中，由串 ABa 可直接派生 $Aaba$，因为 $B\to ab$ 是此文法中的一个产生式。由串 ABa 可派生 $abababa$，因为接连使用产生式 $B\to ab$，$A\to BB$，$B\to ab$ 和 $B\to ab$，可得 $ABa \Rightarrow Aaba \Rightarrow BBaba \Rightarrow Bababa \Rightarrow abababa$。■

定义 4　设 $G=(V, T, S, P)$是一个短语结构文法，由 G 生成的语言（或 G 的语言）是初始符 S 能够派生的所有终结符串构成的集合，记为 $L(G)$。即：

$$L(G)=\{w \in T^* \mid S \stackrel{*}{\Rightarrow} w\}$$

下面两个例子都是寻找短语结构文法所生成的语言。

例 3　设 G 是一个文法，其词汇表为 $V=\{S, A, a, b\}$，终结符集 $T=\{a, b\}$，初始符为 S，产生式为 $P=\{S\to aA, S\to b, A\to aa\}$。求这个文法的语言 $L(G)$。

解　使用产生式 $S\to aA$，可以从初始符 S 派生 aA，还可用产生式 $S\to b$ 派生 b。使用产生式 $A\to aa$，可以从 aA 派生 aaa。没有其他的词还能被派生，故 $L(G)=\{b, aaa\}$。■

例 4　设 G 是一个文法，其词汇表为 $V=\{S, 0, 1\}$，终结符集 $T=\{0, 1\}$，初始符为 S，产生式为 $P=\{S\to 11S, S\to 0\}$。求这个文法的语言 $L(G)$。

解　分别使用 $S\to 0$ 和 $S\to 11S$，可以从 S 派生出 0 和 $11S$。从 $11S$ 可以派生出 110 和 $1111S$。从 $1111S$ 可以派生出 11110 和 $111111S$。在派生过程的每一步，或者在串的末尾加两个 1，或者在串的末尾加 0 后终止派生。总之，$L(G)=\{0, 110, 11110, 111110, \cdots\}$，即 $L(G)$是如下串的集合：开始是偶数个 1，最后是一个 0。这个结论可如下归纳证明：使用 n 次产生式之后，所生成的终结符串只能是这样的串：先是 $n-1$ 个或为数更少的 11 的连接，后面跟一个 0（留作练习）。■

经常出现的问题是要构造一个文法来生成一个给定的语言。下面的三个例子描述这种问题。

例 5 给出生成集合$\{0^n1^n \mid n=0, 1, 2, \cdots\}$的一个短语结构文法。

解 此集合中的元素是这样的串：先是一串 0，后跟含同样多个 1 的串。可以用两个产生式来生成所有这些串(包括空串)，第一个产生式对语言中的串不断地产生更长的串，方法是在串前面加一个 0，串末尾加一个 1；第二个产生式以空串来替代 S。所求的文法是 $G=(V, T, S, P)$，其中 $V=\{0, 1, S\}$，终结符集 $T=\{0, 1\}$，初始符为 S，产生式为

$S\rightarrow 0S1$

$S\rightarrow \lambda$

此文法能够生成所给集合的证明作为练习留给读者。 ■

上面这个例子讨论的是如下串的集合：前面是一串 0，后面跟一串 1，其中 0 的个数和 1 的个数相同。下面的例子还是讨论这样的串，但 0 的个数与 1 的个数不一定相同。

例 6 给出生成集合$\{0^m1^n \mid m$ 和 n 为非负整数$\}$的一个短语结构文法。

解 下面构造生成这个集合的两个文法 G_1 和 G_2。这也说明两个文法可能生成相同的语言。

文法 G_1 的字母表 $V=\{S, 0, 1\}$，终结符集 $T=\{0, 1\}$，产生式为 $S\rightarrow 0S$，$S\rightarrow S1$ 和 $S\rightarrow\lambda$。G_1 能生成所给集合，因为应用第一个产生式 m 次就在串的前面增加了 m 个 0，应用第二个产生式 n 次就在串的后增加了 n 个 1。详细证明留作练习。

文法 G_2 的字母表 $V=\{S, V, 0, 1\}$，终结符集 $T=\{0, 1\}$，产生式为 $S\rightarrow 0S$，$S\rightarrow 1A$，$S\rightarrow 1$，$A\rightarrow 1A$，$A\rightarrow 1$ 和 $S\rightarrow\lambda$。G_2 也能生成所给集合的详细证明留作练习。 ■

有时候，一些很容易描述的集合不得不用非常复杂的文法来生成，下面就是一个这样的例子。

例 7 生成集合$\{0^n1^n2^n \mid n=0, 1, 2, 3, \cdots\}$的一个文法是：$G=\{V, T, S, P\}$，其中 $V=\{0, 1, 2, S, A, B\}$，终结符集 $T=\{0, 1, 2\}$，初始符为 S，产生式有 $S\rightarrow 0SAB$，$S\rightarrow\lambda$，$BA\rightarrow AB$，$0A\rightarrow 01$，$1A\rightarrow 11$，$1B\rightarrow 12$，$2B\rightarrow 22$。此命题的正确性证明留作练习。在本节后面指出的意义下，会明白此文法是生成这个语言的最简单类型的文法。读者也许奇怪这个文法是怎么得来的，因为似乎很难凭空想出这样一个文法。但如果知道可以用计算理论中的技术系统地构造出这个文法，也许就不奇怪了，但这种技术已超出了本书的范围。 ■

11.1.3 短语结构文法的类型

短语结构文法可以根据其产生式的类型来分类。下面我们来描述诺姆·乔姆斯基[⊖]引入的分类方法。在 11.4 节将会看到，以这种方法定义的不同语言类型，与不同的计算机器模型识别的语言类相对应。

0 型文法对其产生式没有限制。1 型文法只有两种形式的产生式：一种是 $w_1\rightarrow w_2$ 形式的产生式，其中 w_2 的长度大于或等于 w_1 的长度；另一种是 $w_1\rightarrow\lambda$。2 型文法只有形如 $w_1\rightarrow w_2$ 的产生式，其中 w_1 是一个非终结符的单个符号。3 型文法只有形如 $w_1\rightarrow w_2$ 的产生式，同时满足 $w_1=A$，且 $w_2=aB$ 或 $w_2=a$，其中 A 和 B 是非终结符，a 是终结符，或者满足 $w_1=S$，$w_2=\lambda$。

⊖ **诺姆·乔姆斯基(Avram Noam Chomsky，生于 1928 年)** 乔姆斯基出生在费城。他的父亲是一位希伯来语的学者。乔姆斯基在宾夕法尼亚大学获得语言学学士、硕士和博士学位。1950 年至 1951 年，他在宾夕法尼亚大学任教。1955 年受聘于麻省理工学院，开始执教法语和德语的生涯。乔姆斯基现今领有麻省理工学院外国语和语言学的费拉雷·华德教授衔。他以在语言学方面的基础贡献而著称，其中包括对文法的研究。乔姆斯基还以对政治的直言不讳而知名于世。

从定义可以看出，3 型文法都是 2 型文法，2 型文法都是 1 型文法，1 型文法都是 0 型文法。2 型文法又称为上下文无关文法，因为如果某个串中出现了一个非终结符，而此非终结符又是某个产生式的左边，则不管此串中的其他符号是什么，这个非终结符都可被换掉。2 型文法生成的语言称为上下文无关语言。当一个文法具有形如 $lw_1r \to lw_2r$（而不是形如 $w_1 \to w_2$）的产生式时，这样的文法称为 1 型文法或上下文相关文法，因为只有当 w_1 被串 l 和 r 包围时，才能替换为 w_2。3 型文法又称为正则文法。正则文法生成的语言称为是正则的。11.4 节讨论正则语言和有限状态机之间的关系。图 11-1 的文氏图说明了这些不同类型文法间的关系。

例 8 由例 6 可知，$\{0^m1^n \mid m, n = 0, 1, 2, \cdots\}$ 是一个正则语言，因为它是由一个正则文法生成的，即由例 6 的文法 G_2 生成的。■

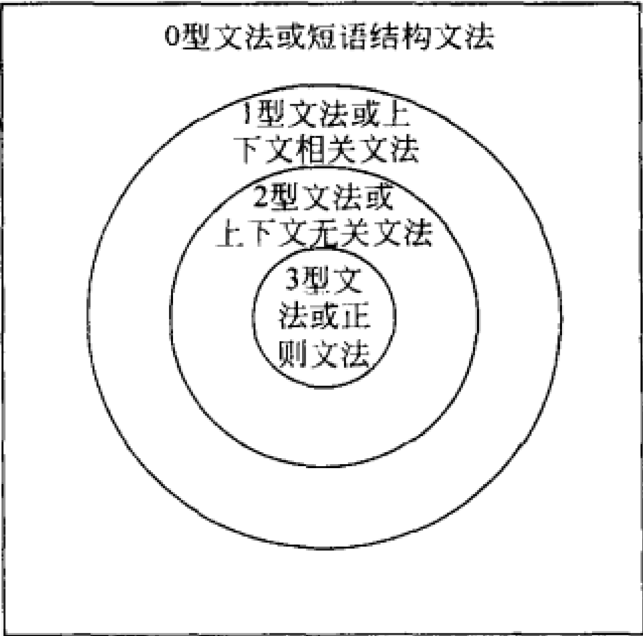


图 11-1 文法的类型

上下文无关文法和正则文法在编程语言中起着重要的作用。上下文无关文法被用于定义几乎所用编程语言的语法，这些语法强得足以定义大多数的语言。而且，可以设计出有效的算法来确定是否以及如何生成一个串。正则文法则用于搜索特定模式的文本和进行词汇分析，词汇分析过程将输入流转变为标记流以供语法分析器使用。

例 9 由例 5 可知，$\{0^n1^n \mid n=0, 1, 2, \cdots\}$ 是一个上下文无关语言，因为这个文法的产生式为 $S \to 0S1$ 和 $S \to \lambda$。在 11.4 节中我们将证明它不是正则语言。■

例 10 集合 $\{0^n1^n2^n \mid n=0, 1, 2, 3, \cdots\}$ 是一个上下文相关语言，因为它是由例 7 中的 1 型文法生成的。但它不是 2 型语言（如本章补充练习中的练习 28 所证）。■

表 11-1 概括了用来对短语结构文法进行分类的术语。

表 11-1 文法的类型

类型	对产生式 $w_1 \to w_2$ 的限制	类型	对产生式 $w_1 \to w_2$ 的限制
0	无限制	2	$w_1 = A$ 其中 A 是非终结符
1	$l(w_1) \leq l(w_2)$，或 $w_2 = \lambda$	3	$w_1 = A$ 和 $w_2 = aB$ 或 $w_2 = a$，其中 $A \in N$，$B \in N$ 和 $a \in T$，或 $S \to \lambda$

11.1.4 派生树

对上下文无关文法生成的语言，其派生可以用有序根树表示成图形，这样的树称为派生树或语法分析树，树根表示初始符；树的内部节点表示在派生过程中产生的非终结符，树的叶节点表示终结符。如果在派生过程中，用到了产生式 $A \to w$（其中 w 是一个词），则表示 A 的节点就有一些子节点，它们表示 w 中的每一个符号，并且从左到右排列。

例 11 对于本节引言所给的例子，构造派生 *the hungry rabbit eats quickly* 的派生树。

解 所求派生树如图 11-2 所示。■

在许多应用中，都会遇到这样的问题：确定一个串是否在一个上下文无关文法生成的语言中，例如编译器的构造。下面例子指出了解决这样问题的两个方法。

例 12 确定词 $cbab$ 是否在文法 $G=\{V, T, S, P\}$ 生成的语言中，其中，$V=\{a, b, c, A, B, C, S\}$，$T=\{a, b, c\}$，$S$ 为初始符，产生式为

$S \to AB$

$A \to Ca$

$B \to Ba$

$B \to Cb$

$B \to b$

$C \to cb$

$C \to b$

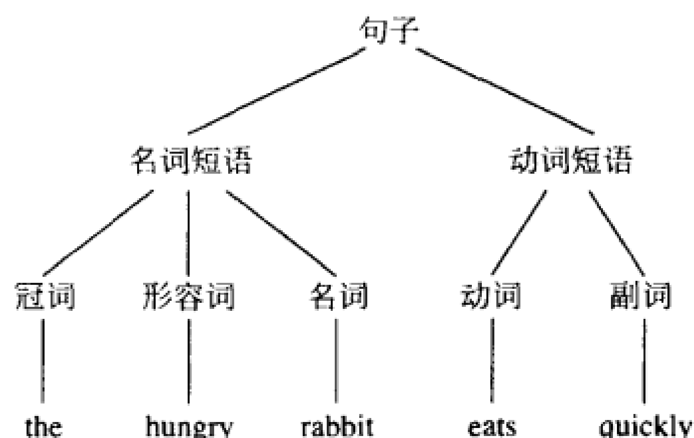


图 11-2 派生树

解 解决这个问题的一种办法是：从 S 出发，用一系列产生式试着派生出 $cbab$。因为只有一个产生式的左边是 S，故必须从 $S \Rightarrow AB$ 开始。下一步，用左边是 A 的唯一产生式 $A \to Ca$ 得到 $S \Rightarrow AB \Rightarrow CaB$。因为 $cbab$ 以符号 cb 开始，故我们使用产生式 $C \to cb$，这样就得到了 $S \Rightarrow AB \Rightarrow CaB \Rightarrow cbaB$。最后，使用产生式 $B \to b$ 就可得到 $S \Rightarrow AB \Rightarrow CaB \Rightarrow cbaB \Rightarrow cbab$。这种方法称为自顶向下的语法分析，因为它从初始符号开始，一个接一个地用产生式来处理。

解决整个问题的另一个办法称为自底向上的语法分析，这种办法从后向前处理。因为 $cbab$ 是需要派生的串，故可以使用产生式 $C \to cb$，从而得到 $Cab \Rightarrow cbab$。再使用产生式 $A \to Ca$ 得到 $Ab \Rightarrow Cab \Rightarrow cbab$。由产生式 $B \to b$ 可得 $AB \Rightarrow Ab \Rightarrow Cab \Rightarrow cbab$。最后再用产生式 $S \to AB$，就可得到 $cbab$ 的一个完整的派生 $S \Rightarrow AB \Rightarrow Ab \Rightarrow Cab \Rightarrow cbab$。■

11.1.5 巴克斯-诺尔范式

有时候还用另一个方法来表示 2 型文法，这就是巴克斯-诺尔范式，这个方法是根据约翰·巴克斯[⊖]和彼得·诺尔[⊜]命名的，约翰·巴克斯是它的发明人，彼得·诺尔则改进了它，并将之应用于程序设计语言 ALGOL 的规范说明。巴克斯-诺尔范式已用来对许多程序设计语言(包括 Java)的语法规则进行规范说明。在 2 型文法中，产生式的左边都是单个非终结符。在巴克斯-诺尔范式中，将左边是同一个非终结符的所有产生式合并成一个式子，而不是将

⊖ **约翰·巴克斯(John Backus，生于 1924 年)** 巴克斯生于费城，在哥伦比亚大学获得理学学士和数学硕士学位。于 1950 年进入 IBM 当程序设计员。他参加了 IBM 的两种早期计算机的设计与开发。从 1954 年到 1958 年，他领导 IBM 的一个小组开发了 FORTRAN。巴克斯在 1958 年成为 IBM 沃森研究中心的职员。他是程序设计语言 ALGOL 设计委员会的一员。正是在此语言的设计过程中，他使用了现今称为巴克斯-诺尔范式的方法来描述此语言的语法。后来，巴克斯从事于集合簇的数学研究和函数型程序设计的研究。巴克斯于 1963 年成为 IBM 的特别会员(fellow)，于 1974 年获美国国家科学奖(National Medal of Science)，于 1977 年获美国计算机协会(Association of Computing Machinery)颁发的具有崇高声誉的图灵奖。

⊜ **彼得·诺尔(Peter Naur，生于 1928 年)** 诺尔生于哥本哈根附近的 Frederiksberg。孩提时代，诺尔就对天文学感兴趣。他不局限于观察天体，还计算彗星和小行星的轨道。诺尔毕业于哥本哈根大学，并于 1949 年获得学位。他在剑桥度过了 1950 和 1951 年，在此期间他用早期的计算机来计算彗星和行星的运动。回到丹麦后，他虽然继续从事天文学的研究，但保持着与计算的紧密联系。在 1955 年，他作为顾问参与丹麦的第一台计算机的建造。在 1959 年，诺尔从天文学转入计算的研究，并将计算的研究作为专职工作。作为一个专职计算机科学家，他的第一个工作是参加程序设计语言 ALGOL 的开发。从 1960 至 1967 年，他继续从事 ALGOL 和 COBOL 编译器的研究。1969 年，他成为哥本哈根大学的计算机科学教授，在程序设计方法学领域辛勤耕耘着。

这些产生式都列出来。我们还用符号：：＝代替→，将非终结符用<>括起来，并在一个式子里列出所有这些产生式的右边，用竖线将这些产生式分开。例如，产生式 $A\rightarrow Aa$，$A\rightarrow a$，$A\rightarrow AB$ 可以合并成〈A〉::＝〈A〉a | a | 〈A〉〈B〉。

例 13 给出了如何用巴克斯-诺尔范式来描述编程语言的语法。本例来自于巴克斯-诺尔范式在 ALGOL 60 中的使用。

例 13 在 ALGOL 中，标识符(如同变量这样的实体的名字)由字母数字字符串组成，且必须以字母打头。我们可以用巴克斯-诺尔范式描述可用的标识符集合。

〈标识符〉::＝〈字母〉|〈标识符〉〈字母〉|〈标识符〉〈数字〉

〈字母〉::＝*a* | *b* | … | *y* | *z*(省略号表示包括全部 26 个字母)

〈数字〉::＝0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9

例如，我们可以如下生成有效的标识符 *x*99*a*：用第一条规则将〈标识符〉替换成〈标识符〉〈字母〉，用第二条规则得到〈标识符〉*a*，两次使用第一条规则得到〈标识符〉〈数字〉〈数字〉*a*，两次使用第三条规则得到〈标识符〉99*a*，再用第一条规则得到〈字母〉99*a*，最后用第二条规则得到 *x*99*a*。 ■

例 14 本节引言描述了英语的一个子集，其对应文法的巴克斯-诺尔范式是什么？

解 这个文法的巴克斯-诺尔范式是：

〈句子〉::＝〈名词短语〉〈动词短语〉

〈名词短语〉::＝〈冠词〉〈形容词〉〈名词〉|〈冠词〉〈名词〉

〈动词短语〉::＝〈动词〉〈副词〉|〈动词〉

〈冠词〉::＝*a* | *the*

〈形容词〉::＝*large* | *hungry*

〈名词〉::＝*rabbit* | *mathematician*

〈动词〉::＝*eats* | *hops*

〈副词〉::＝*quickly* | *wildly* ■

例 15 给出带符号十进制整数的产生式的巴克斯-诺尔范式(带符号整数是非负整数前面加上一个加号或减号)。

解 一个产生带符号整数的文法之巴克斯-诺尔范式为：

〈带符号整数〉::＝〈符号〉〈整数〉

〈符号〉::＝＋ | －

〈整数〉::＝〈数字〉|〈数字〉〈整数〉

〈数字〉::＝0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 ■

具有多种扩展的巴克斯-诺尔范式广泛用于说明编程语言(如 Java 和 LISP)、数据库语言(如 SQL)和标记语言(如 XML)的语法。一些在编程语言描述中常用的巴克斯-诺尔范式在本节末练习 28 的导语中有介绍。

练习

练习 1～3 中的文法是：初始符为句子，终结符集 T＝{*the*，*sleepy*，*happy*，*tortoise*，*hare*，*passes*，*runs*，*quickly*，*slowly*}，非终结符集 N＝{名词短语，及物动词短语，不及物动词短语，冠词，形容词，名词，动词，副词}，产生式为

句子→名词短语 及物动词短语 名词短语

句子→名词短语 不及物动词短语

名词短语→冠词　形容词　名词

名词短语→冠词　名词

及物动词短语→及物动词

不及物动词短语→不及物动词　副词

不及物动词短语→不及物动词

冠词→*the*

形容词→*sleepy*

形容词→*happy*

名词→*tortoise*

名词→*hare*

及物动词→*passes*

不及物动词→*runs*

副词→*quickly*

副词→*slowly*

1. 用产生式集证明下列每个句子都是有效句子：

a) *the happy hare runs*

b) *the sleepy tortoise runs quickly*

c) *the tortoise passes the hare*

d) *the sleepy hare passes the happy tortoise*

2. 除了练习 1 中的有效句子外，再给出五个有效句子。

3. 证明：*the hare runs the sleepy tortoise* 不是有效句子。

*4. 设 $V=\{S, A, B, a, b\}$，$T=\{a, b\}$。当产生式集为下列情形之一时，求文法 $\{V, T, S, P\}$ 生成的语言。

a) $S\to AB$，$A\to ab$，$B\to bb$。

b) $S\to AB$，$S\to aA$，$A\to a$，$B\to ba$。

c) $S\to AB$，$S\to AA$，$A\to aB$，$A\to ab$，$B\to b$。

d) $S\to AA$，$S\to B$，$A\to aaA$，$A\to aa$，$B\to bB$，$B\to b$。

e) $S\to AB$，$A\to aAb$，$B\to bBa$，$A\to\lambda$，$B\to\lambda$。

5. 用例 5 所给的文法构造 $0^3 1^3$ 的派生。

6. 证明：例 5 所给的文法生成集合 $\{0^n 1^n \mid n=0, 1, 2, \cdots\}$。

7. a) 用例 6 中的文法 G_1 构造 $0^2 1^4$ 的派生。

b) 用例 6 中的文法 G_2 构造 $0^2 1^4$ 的派生。

8. a) 证明：例 6 中的文法 G_1 生成集合 $\{0^m 1^n \mid m, n=0, 1, 2, \cdots\}$。

b) 证明：例 6 中的文法 G_2 生成同一个集合。

9. 用例 7 所给的文法构造 $0^2 1^2 2^2$ 的派生。

*10. 证明：例 7 所给的文法生成集合 $\{0^n 1^n 2^n \mid n=0, 1, 2, \cdots\}$。

*11. 求下列语言的短语结构文法：

a) 包含偶数个 0 但没有 1 的所有二进制串构成的集合。

b) 由 1 后面跟奇数个 0 的所有二进制串构成的集合。

c) 包含偶数个 0 和偶数个 1 的所有二进制串构成的集合。

d) 包含 10 个以上 0 但没有 1 的所有二进制串构成的集合。

e) 所包含 0 的个数多于 1 的个数的所有二进制串构成的集合。

f) 包含相同个数的 0 和 1 的所有二进制串构成的集合。

g) 包含不同个数的 0 和 1 的所有二进制串构成的集合。

12. 构造生成下列集合的短语结构文法：

a) $\{01^{2n} \mid n\geqslant 0\}$。　b) $\{0^n 1^{2n} \mid n\geqslant 0\}$。　c) $\{0^n 1^m 0^n \mid m\geqslant 0, n\geqslant 0\}$。

13. 设 $V=\{S, A, B, a, b\}$，$T=\{a, b\}$。若产生式集 P 为下列集合时，问文法 $G=(V, T, S, P)$ 是否

为 0 型但不是 1 型文法？是否为 1 型但不是 2 型文法？或是否为 2 型但不是 3 型文法？

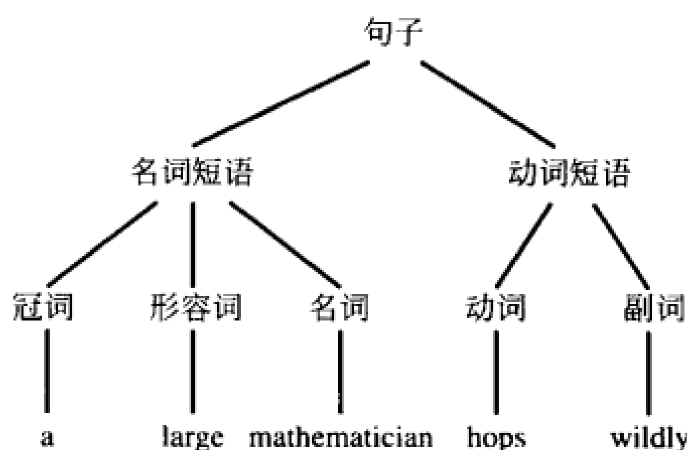
a) $S\to aAB$，$A\to Bb$，$B\to\lambda$。
b) $S\to aA$，$A\to a$，$A\to b$。
c) $S\to ABa$，$AB\to a$。
d) $S\to ABA$，$A\to aB$，$B\to ab$。
e) $S\to bA$，$A\to B$，$B\to a$。
f) $S\to aA$，$aA\to B$，$B\to aA$，$A\to b$。
g) $S\to bA$，$A\to b$，$S\to\lambda$。
h) $S\to AB$，$B\to aAb$，$aAb\to b$。
i) $S\to aA$，$A\to bB$，$B\to b$，$B\to\lambda$。
j) $S\to A$，$A\to B$，$B\to\lambda$。

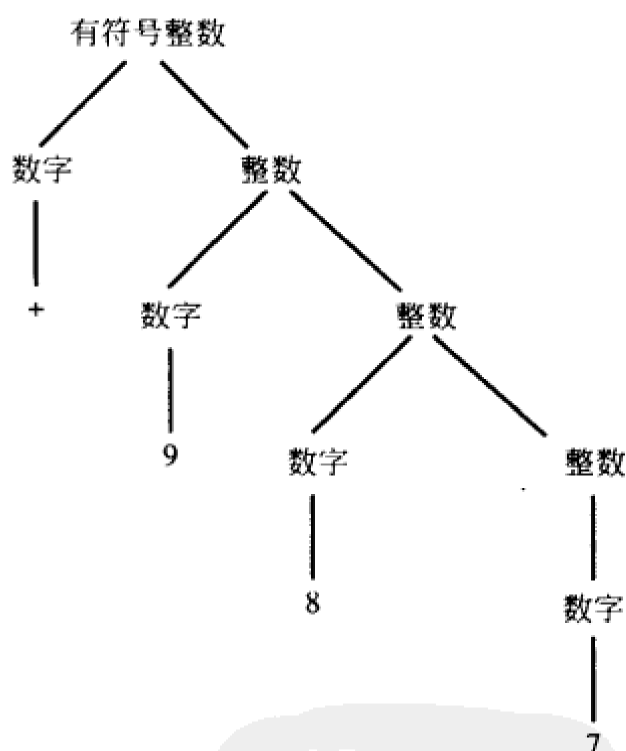
14. 回文是从前向后读和从后向前读都一样的串，也就是 $w=w^R$ 这样的串 w，其中 w^R 是串 w 的倒转。试求一个上下文无关的文法，使得其生成的集合是字母表{0，1}上的所有回文。

*15. 设 G_1 和 G_2 是两个上下文无关的文法，它们生成的语言分别为 $L(G_1)$ 和 $L(G_2)$。试证：对于下列每个集合，都有一个上下文无关文法来生成下列集合：

a) $L(G_1)\cup L(G_2)$　　b) $L(G_1)L(G_2)$　　c) $L(G_1)^*$

16. 求用右面的两幅图派生树构造的串。





17. 构造练习 1 中的句子的派生树。

18. 设 G 是一个文法，其中 $V=\{a, b, c, S\}$，$T=\{a, b, c\}$，初始符号为 S，产生式为 $S\to abS$，$S\to bcS$，$S\to bbS$，$S\to a$，$S\to cb$。构造下列串的派生树。

a) $bcbba$　　b) $bbbcbba$　　c) $bcabbbbbcb$

*19. 对于下列每个串，用自顶向下的语法分析方法，确定其是否属于例 12 中的文法生成的语言。

a) $baba$　　b) $abab$　　c) $cbaba$　　d) $bbbcba$

*20. 对于练习 19 中的串，用自底向上的语法分析方法，确定其是否属于例 12 中的文法生成的语言。

21. 用例 15 所给的文法构造－109 的派生树。

22. a) 如果一个文法的产生式由下列巴克斯-诺尔范式给出，这些产生式是什么？

〈表达式〉∷=(〈表达式〉)|〈表达式〉+〈表达式〉|〈表达式〉*〈表达式〉|〈变元〉

〈变元〉∷=x|y

b) 求此文法中 $(x*y)+x$ 的派生树。

23. a) 构造一个短语结构文法，使其生成如下所有带符号十进制数：这些数由符号(＋或－)、非负整数和十进制小数三部分构成，且十进制小数部分或者是空串，或者是小数点后面跟一个正整数，其中，整数的开始部分允许有一些 0。

b) 给出这个文法的巴克斯-诺尔范式。　　c) 构造此文法中－31.4 的派生树。

24. a) 构造一个短语结构文法，使其生成所有形如 a/b 的分数构成的集合，其中 a 为带符号十进制数，b 是正整数。

b) 给出这个文法的巴克斯-诺尔范式。　　c) 构造此文法中＋311/17 的派生树。

25. 对于包含如下内容的标识符，给出其巴克斯-诺尔范式的产生式规则：

a) 一个或多个小写字母。　　b) 至少三个但至多六个小写字母。

c) 一至六个大写或小写字母并以大写字母开头。

d) 一个小写字母，后跟一个数字或下划线，后跟三或四个字母数字字符(大小写字母和数字)。

26. 给出如下人名的巴克斯-诺尔范式产生式规则，人名包含："名"，它是一个仅有首字母大写的字母串；"中间名"；"姓"，它可以是任意字母串。

27. 给出C语言中生成所有标识符的巴克斯-诺尔范式产生式规则。在C语言中，标识符以一个字母或下划线开始，后跟一或多个小写字母、大写字母、下划线和数字。

巴克斯-诺尔范式的一些扩展常用于定义短语结构文法。在其中的一种扩展中，问号(?)表明其左边的符号(或括号中的一组符号)可以出现零次或一次(即它是可选的)，星号(*)表明其左边的符号可出现零次或多次，加号(+)表明其左边的符号可出现一次或多次。这些扩展均为**扩展的巴克斯-诺尔范式**(EBNF)的一部分，符号?，*和+称为**元字符**。在EBNF中，用于表示非终结符的括号通常不显示。

28. 描述由下列EBNF产生式集合定义的串的集合。

a) $string ::= L + D?\ L+$
　$L ::= a \mid b \mid c$
　$D ::= 0 \mid 1$

b) $string ::= sign\ D+ \mid D+$
　$sign ::= + \mid -$
　$D ::= 0 \mid 1 \mid 2 \mid 3 \mid 4 \mid 5 \mid 6 \mid 7 \mid 8 \mid 9$

c) $string ::= L * (D+)?\ L *$
　$L ::= x \mid y$
　$D ::= 0 \mid 1$

29. 给出生成下述十进制数的扩展的巴克斯-诺尔范式的产生式规则：此十进制数由可选的符号、非负整数和小数部分组成，小数部分或者为空串，或者为小数点后加一个可选的正整数，这个正整数前可能带有若干个零。
30. 如果三明治由下列东西组成：底部的一片面包；芥末或蛋黄酱；生菜(可选)；一片西红柿(可选)；一片或多片火鸡肉、鸡肉或烤牛肉(任意组合)；一些奶酪(可选)；顶部的一片面包。给出其扩展的巴克斯-诺尔范式的产生式规则。
31. 给出C语言中的标识符的EBNF产生式规则(见练习27)。
32. 描述如何将文法的EBNF产生式翻译成巴克斯-诺尔范式产生式。

下面给出的是在后缀(或逆波兰)记法中描述表达式的语法的巴克斯-诺尔范式。

$\langle expression\rangle ::= \langle term\rangle \mid \langle term\rangle\langle term\rangle\langle addOperator\rangle$
$\langle addOperator\rangle ::= + \mid -$
$\langle term\rangle ::= \langle factor\rangle \mid \langle factor\rangle\langle factor\rangle\langle mulOperator\rangle$
$\langle mulOperator\rangle ::= * \mid /$
$\langle factor\rangle ::= \langle identifier\rangle \mid \langle expression\rangle$
$\langle identifier\rangle ::= a \mid b \mid \cdots \mid z$

33. 对下列字符串，判断其是否由后缀记法的文法生成。如果是，给出生成步骤。
a) $abc*+$　b) $xy++$　c) $xy-z*$　d) $wxyz-*/$　e) $ade-*$
34. 用巴克斯-诺尔范式描述中缀记法中表达式的语法，其中运算符和标识符与练习33前的导言中后缀表达式的BNF相同，但对用作因子的表达式必须加括号。
35. 对下列字符串，判断其是否由练习34中的中缀表达式文法生成。若是，给出生成步骤。
a) $x+y+z$　b) $a/b+c/d$　c) $m*(n+p)$　d) $+m-n+p-q$　e) $(m+n)*(p-q)$
36. 设G是一个文法，R是一个关系，有序对$(w_0, w_1)\in R$当且仅当w_1可以从w_0在G中直接派生出来。求R的自反传递闭包。

11.2 带输出的有限状态机

11.2.1 引言

许多种类的机器，包括计算机的某些部件，都可以用有限状态机作为模型。经常用来作为模型的有限状态机也有多种形式，但所有这些形式都包括一个有限的状态集合(其中有一个指定的初始状态)、一个输入字母表和一个转移函数(对每个由状态和输入构成的对指定下一个状态)。有限状态机广泛应用于计算机科学和数据网络中。例如，有限状态机是许多程序的基础，如拼写检查、语法检查、索引或搜索大的文本、语音识别、采用如XML和HTML标

记语言转换文本，以及规范计算机如何通信的网络协议。

本节将介绍有限状态机产生的结果。如，有限状态机如何用来模拟自动售货机、输入延迟机、整数加法器以及判断位串是否包含指定模式的机器。

在给出形式定义之前，先来说明怎么建立自动售货机的模型。自动售货机可以接受 5 分、1 角和 25 分硬币。如果 30 分或更多硬币被投到机器里，则机器立刻退出超过 30 分的部分。如果顾客投放了 30 分且超出部分已被退还，则顾客可以按橙色按钮得到一筒橘子汁，或者按红色按钮得到一筒苹果汁。可以如下描述这个机器是怎么工作的：详细描述它的状态，且说明它在接受输入后怎么改变状态，还要说明对输入和当前状态的各种组合，它所产生的输出。

这个机器可能处于 7 种状态 $s_i(i=0, 1, 2, \cdots, 6)$，其中状态 s_i 指机器已经收集了 $5i$ 分。机器以表示收集了 0 分的状态 s_0 开始。输入可能是：5 分、1 角、25 分、橙色钮(O)或红色钮(R)。输出可能是：空(n)、5 分、1 角、15 分、20 分、25 分、一筒橘子汁或一筒苹果汁。

下面的例子说明了此机器的模型是怎么工作的。假设一个学生先投入了 1 角，又投入了 25 分，得到了 5 分的找赎，然后按橙色按钮就得到一筒橘子汁。机器从状态 s_0 开始。它的第一个输入是 10 分，这就将机器的状态改变为 s_2，但没有输出。第二个输入是 25 分，这将状态从 s_2 改变为 s_6，并返回 5 分作为输出。下一个输入是橙色按钮，它将状态从 s_6 改回到 s_0(因为机器返回到初始状态)，并送出一筒橘子汁作为输出。

可以将机器的所有这些状态变化和输出用一个表来表示。为此，对状态和输入的每个组合，我们都需要指明下一个状态和产生的输出。表 11-2 对每对状态和输入都指明了转移和输出。

表 11-2 自动售货机的状态表

状态	下一个状态					输出				
	输入					输入				
	5	10	25	O	R	5	10	25	O	R
s_0	s_1	s_2	s_5	s_0	s_0	n	n	n	n	n
s_1	s_2	s_3	s_6	s_1	s_1	n	n	n	n	n
s_2	s_3	s_4	s_6	s_2	s_2	n	n	5	n	n
s_3	s_4	s_5	s_6	s_3	s_3	n	n	10	n	n
s_4	s_5	s_6	s_6	s_4	s_4	n	n	15	n	n
s_5	s_6	s_6	s_6	s_5	s_5	n	5	20	n	n
s_6	s_6	s_6	s_6	s_0	s_0	5	10	25	OJ	AJ

说明机器动作的另一个方法是使用带标号边的有向图，其中：状态表示为小圈，边表示转移，并用输入和这个转移产生的输出对边进行标号。自动售货机的有向图如图 11-3 所示。

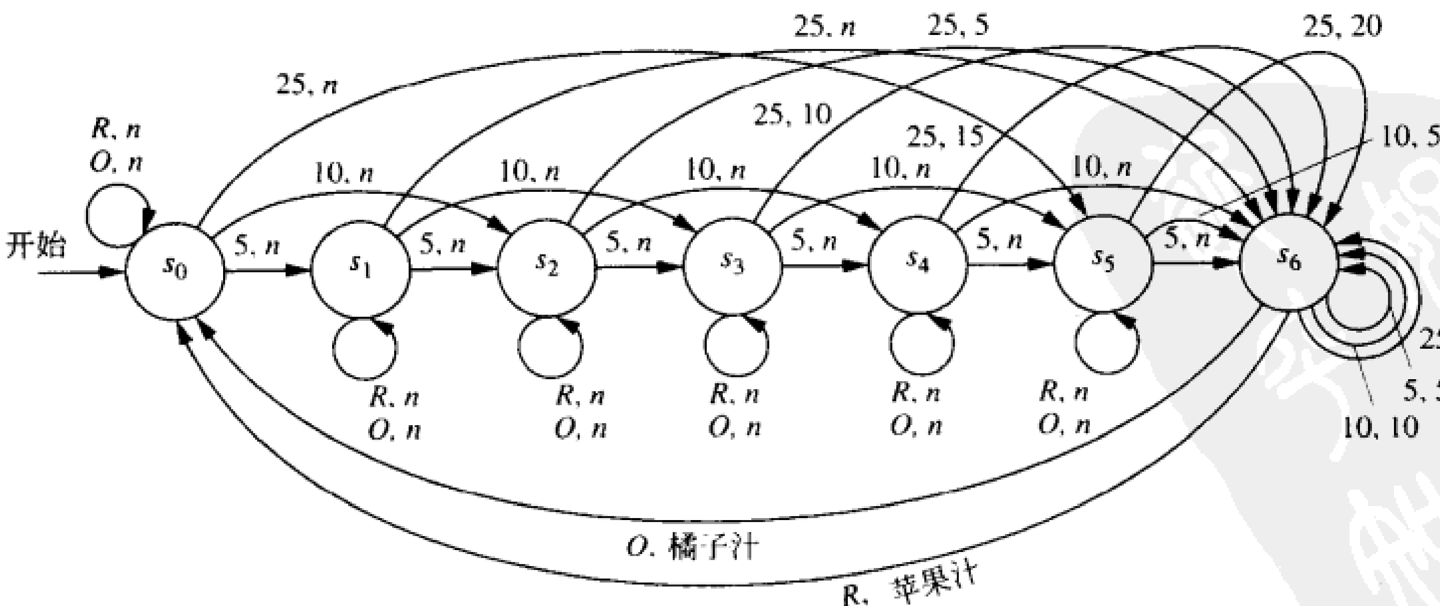


图 11-3 一个自动售货机

11.2.2 带输出的有限状态机

现在给出带输出的有限状态机的形式定义。

定义 1 有限状态机 $M=(S, I, O, f, g, s_0)$ 由如下部分组成：一个有限的状态集合 S，一个有限的输入字母表 I，一个有限的输出字母表 O，一个转移函数 f，f 为每个状态和输入对指派一个新状态，一个输出函数 g，g 为每个状态和输入对指派一个输出，还有一个初始状态 s_0。

设 $M=(S, I, O, f, g, s_0)$ 是一个有限状态机，可以用状态表来表示状态函数和输出函数的值。在本节引言中，我们已经构造了自动售货机的状态表。

例 1 表 11-3 的状态表描述了一个有限状态机，其中 $S=\{s_0, s_1, s_2, s_3\}$，$I=\{0, 1\}$，且 $O=\{0, 1\}$。转移函数 f 的值在前两列给出，输出函数 g 的值在后两列给出。■

表示有限状态机的另一种方法是用状态图，这是一个带标号边的有向图。在这个图中，状态由圈表示，转移由带标号的输入和输出对箭头表示。

例 2 构造状态表如表 11-3 所示的有限状态机的状态图。

解 这个机器的状态图如图 11-4 所示。■

表 11-3

状态	f 输入 0	f 输入 1	g 输入 0	g 输入 1
s_0	s_1	s_0	1	0
s_1	s_3	s_0	1	1
s_2	s_1	s_2	0	1
s_3	s_2	s_1	0	0

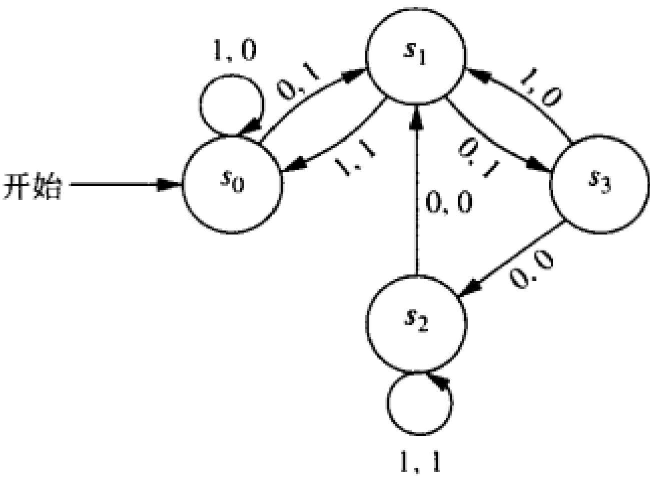


图 11-4 表 11-3 所示有限状态机的状态图

例 3 构造状态图如图 11-5 所示的有限状态机的状态表。

解 这个机器的状态表如表 11-4 所示。■

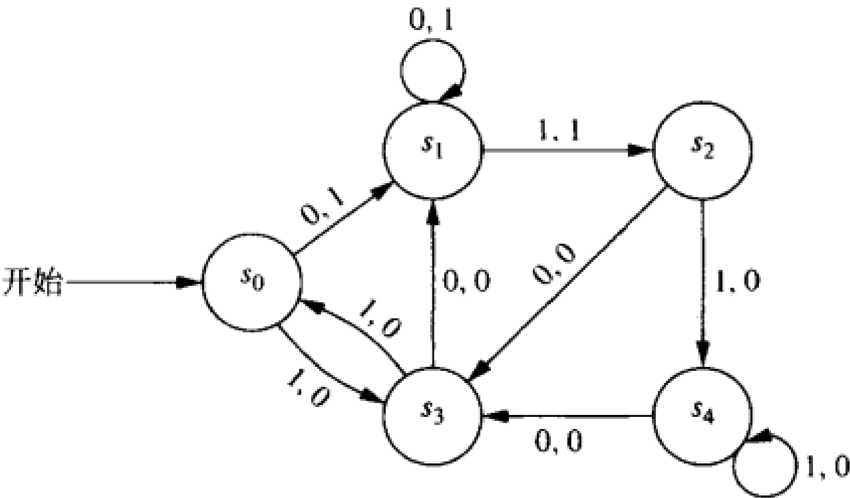


图 11-5 一个有限状态机

表 11-4

状态	f 输入 0	f 输入 1	g 输入 0	g 输入 1
s_0	s_1	s_3	1	0
s_1	s_1	s_2	1	1
s_2	s_3	s_4	0	0
s_3	s_1	s_0	0	0
s_4	s_3	s_4	0	0

一个输入串使初始状态经历一系列状态，这些状态都是由转移函数确定的。当我们（从左向右）一个符号一个符号地读输入串的时候，每个输入符号都使机器从一个状态变为另一个状态。

因为每个转移产生一个输出，故一个输入串产生一个输出串。

设输入串为 $x=x_1x_2\cdots x_k$。则读这个输入使得机器从状态 s_0 变为状态 s_1，其中 $s_1=f(s_0, x_1)$，然后变为状态 s_2，其中 $s_2=f(s_1, x_2)$，等等，以状态 $s_k=f(s_{k-1}, x_k)$结束。这个转移序列就产生了输出串 $y=y_1y_2\cdots y_k$，其中 $y_1=g(s_0, x_1)$是对应于从 s_0 到 s_1 的转移的输出，$y_2=g(s_1, x_2)$是对应于从 s_1 到 s_2 的转移的输出，等等。一般地，$y_j=g(s_{j-1}, x_j)$，$j=1, 2, \cdots, k$。这样我们可以将输出函数 g 的定义扩展到输入串，即定义 $g(x)=y$，其中 y 是对应输入串 x 的输出。在许多应用中，这个记号都很有用。

例 4 对于图 11-5 表示的有限状态机，求其对输入串 101011 生成的输出串。

解 输出是 001000。状态和输出的逐次变化如表 11-5 所示。 ■

表 11-5

输入	1	0	1	0	1	1	—
状态	s_0	s_3	s_1	s_2	s_3	s_0	s_3
输出	0	0	1	0	0	0	—

我们现在来看几个有限状态机的有用的例子。例 5、例 6 和例 7 表示了内存容量受限的有限状态机的状态。这些状态用来记住机器读取的符号的属性。然而，由于状态数有限，有限状态机不能用于一些重要的目的。这一点会在 11.4 节讲到。

例 5 单位延迟机是许多电子装置中的一个重要部件，它将输入串延迟一定时间量后输出。怎么构造一个有限状态机使其将输入串延迟一个单位时间呢？即：对于输入的二进位串 $x_1x_2\cdots x_k$，怎么才能输出二进位串 $0x_1x_2\cdots x_{k-1}$？

解 可以如下构造一个延迟机：它有两种可能的输入，即 0 和 1，它还必须有一个初始状态 s_0。因为它还要记住前一个输入是 0 还是 1，所以还需要另外两个状态 s_1 和 s_2，使得如果前一个输入是 1，则机器处于状态 s_1，如果前一个输入是 0，则机器处于状态 s_2。从 s_0 出发的第一个转移产生输出 0，从 s_1 出发的每个转移都产生输出 1，从 s_2 出发的每个转移都产生输出 0。则对应于输入串 $x_1x_2\cdots x_k$ 的输出是这样一个串：以 0 开始，后面跟 x_1，再跟 x_2，…，最后以 x_{k-1} 结束。这个机器的状态图如 11-6 所示。 ■

例 6 试构造一个有限状态机，使其利用整数的二进制展开式将两个整数相加。

解 一般按如下过程将$(x_n\cdots x_1x_0)_2$ 和$(y_n\cdots y_1y_0)_2$ 相加(如 2.5 节所描述)：首先，将数位 x_0 和 y_0 相加，产生和位 z_0 与进位 c_0，且此进位要么是 0，要么是 1；然后将数位 x_1，y_1 连同进位 c_0 一起相加，产生和位 z_1 和进位 c_1；将这个过程一直进行下去；第 n 步是将 x_n，y_n 连同前一个进位 c_{n-1}一起相加，产生和位 z_n 和进位 c_n，c_n 也就是和位 z_{n+1}。

只用两个状态就能构造执行这个加法的有限状态机。为简单起见，假设两个初始位 x_n 和 y_n 都是 0(否则，必须对和位 z_{n+1}作特殊安排)。我们用初始状态 s_0 表示前一个进位是 0(或者是最右边数位的加法)，用另一个状态 s_1 表示前一个进位是 1。因为这个机器的输入是一对二进制数，所以只有四种可能的输入。这四种可能的输入为：00(两个二进制数都是 0)、01(第一个二进制数为 0，第二个为 1)、10(第一个二进制数为 1，第二个为 0)和 11(两个二进制数都是 1)。转移和输出是根据下面两个因素来构造的：一个是输入所表示的两个二进制数的和，另一个是状态所表示的进位。例如：当机器处于状态 s_1 且所接受的输入是 01 时，则下一个状态是 s_1 且输出是 0，因为所产生的和是 $0+1+1=(10)_2$。此机器的状态图如图 11-7 所示。 ■

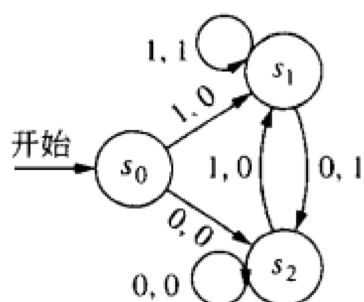


图 11-6 一个单位延迟机

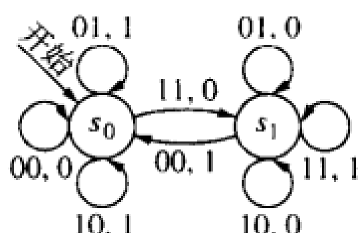


图 11-7 一个做加法的有限状态机器

例 7 在某种编码方法中，当一个信息中出现了 3 个连续的 1，则信息接收器就知道已经发生了一个传送错误。试构造一个有限状态机，使得它输出 1，当且仅当它所接收的最后 3 位都是 1。

解 这个机器需要三个状态。初始状态 s_0 记住前一个输入值(如果存在话)不是 1；状态 s_1 记住前一个输入值是 1，但再前一个输入(如果存在的话)不是 1；状态 s_2 记住前两个输入值都是 1。输入一个 1 将状态 s_0 变为 s_1，因为机器现在读到的是单个的 1，而不是两个连续的 1；它将 s_1 变为 s_2，因为它现在读到了两个连续的 1；它还将 s_2 变为 s_2 本身，因为它已经至少读到了两个连续的 1。输入一个 0 将每个状态都变为 s_0，因为这打断了任何由连续 1 构成的串。如果现在机器所读的是 1，则由 s_2 到 s_2 自身的转移所产生的输出为 1，因为此状态与输入的组合表明机器已经读到了 3 个连续的 1。其他情形的输入都是 0。此机器的状态图如图 11-8 所示。 ■

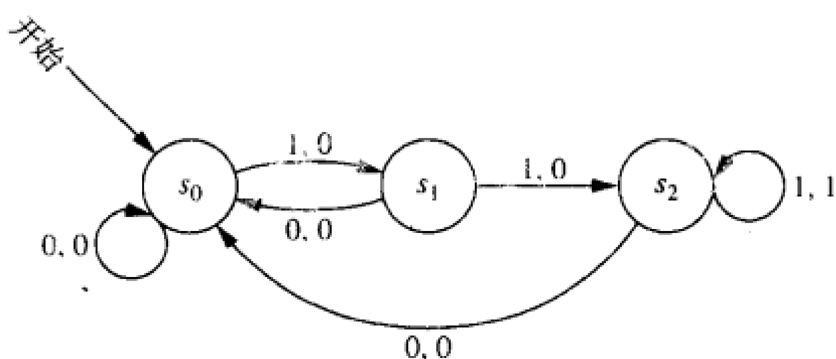


图 11-8 一个有限状态机，它输出 1 当且仅当所读的输入串以 111 结尾

图 11-8 所示的机器是语言识别器的一个例子，因为它输出 1 当且仅当它所读到的输入串具有某种特定的性质。语言识别是有限状态机的一个重要应用。

有限状态机的类型 为建立计算机的模型，人们开发了许多种不同的有限状态机。本节给出了一类有限状态机的定义，在这种类型的机器中，输出与状态间的转移相对应，这种类型的机器称作米利(Mealy)机，因为它是由米利(G. H. Mealy)在 1955 年首先研究的。还有另外一类重要的带输出的有限状态机，其输出仅仅由状态确定，这种类型的有限状态机称为摩尔机，因为它是摩尔(E. F. Moore)在 1956 年提出的。本节结尾有一系列练习讨论摩尔机。

例 7 说明了怎么用米利机来识别语言，但我们通常用另一种不带输出的有限状态机来识别语言。不带输出的有限状态机也称为有限状态自动机，它有一个由终结状态组成的集合，它识别一个串当且仅当能够将初始状态变为一个终结状态。11.3 节将讨论这种类型的有限状态机。

练习

1. 画出具有下列状态表的有限状态机的状态图。

a)

状态	f 输入 0	f 输入 1	g 输入 0	g 输入 1
s_0	s_1	s_0	0	1
s_1	s_0	s_2	0	1
s_2	s_1	s_1	0	0

b)

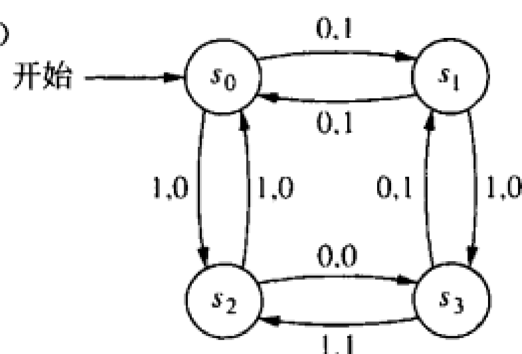
状态	f 输入 0	f 输入 1	g 输入 0	g 输入 1
s_0	s_1	s_0	0	0
s_1	s_2	s_0	1	1
s_2	s_0	s_3	0	1
s_3	s_1	s_2	1	0

c)

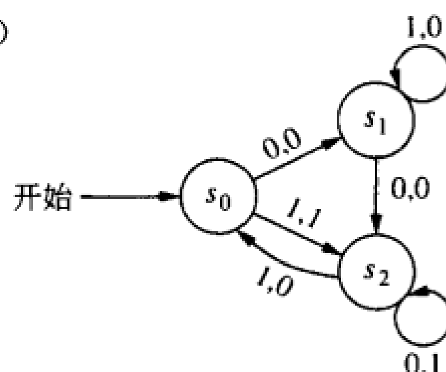
状态	f 输入 0	f 输入 1	g 输入 0	g 输入 1
s_0	s_0	s_4	1	1
s_1	s_0	s_3	0	1
s_2	s_0	s_2	0	0
s_3	s_1	s_1	1	1
s_4	s_1	s_0	1	0

2. 给出具有下列状态图的有限状态机的状态表。

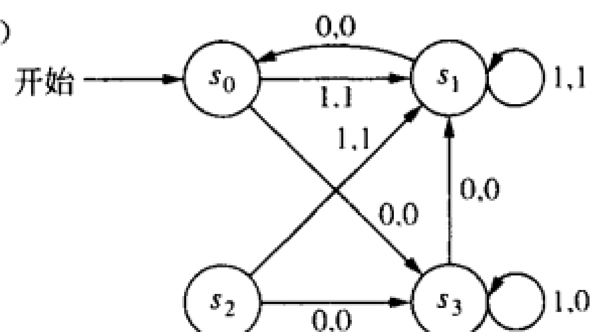
a)



b)



c)



3. 在例 2 所给的有限状态机中，对于下列每个输入串，试确定其输出。

a)0111　　b)11011011　　c)01010101010

4. 在例 3 所给的有限状态机中，对于下列每个输入串，试确定其输出。

a)0000　　b)101010　　c)11011100010

5. 试构造一个有限状态机作为下列饮料机的模型：饮料机接受五分、一角和 25 分的硬币，它一直到接受了 35 分钱币时才开始找回零钱，退出超过 35 分的所有钱币。然后顾客就可以按一些按钮，得到一筒可乐，或一瓶软饮料，或一瓶姜汁啤酒。

6. 试构造一个有限状态机作为下列售报机的模型：它有一个门，此门只在下列两种情形下才可打开：一是放入了 3 个一角硬币(和任意数量的其他硬币)，二是放入了一个 25 分的硬币和一个 5 分的硬币(和任意数量的其他硬币)。一旦门能够被打开，顾客就打开门，取出一份报纸，再关上门。不管塞进去多少额外的钱币，机器都不找回零钱。下一个顾客重新开始时也不能使用上一位多余的钱。

7. 构造一个有限状态机，将输入延迟两位，且以 00 作为输出的头两位。

8. 构造一个有限状态机，对输入串每隔一位改变一次值，且从第二位开始。但保持其他位不变。

9. 构造一个有限状态机来模拟计算机的登录过程：用户首先输入用户标识码，然后输入口令；用户标识码和口令分别被看作是一个输入；如果输入的口令不对，则要求用户重新输入用户标识码。

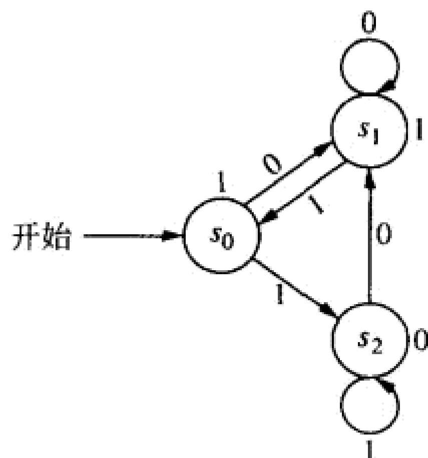
10. 构造一个有限状态机来模拟密码锁，此锁包含数 1 到 40，它只有在输入正确的组合时才能被打开，正确组合是：10 次右，8 次左，37 次右。每个输入都是“一个数、旋转方向、在此方向旋转锁的次数”构成的三元组。

11. 构造一个有限状态机来模拟下列道路收费机：放入 25 分钱币之后此机器将打开一个门。可以使用面额为 5 分、一角和 25 分的硬币。不找零钱，多于 25 分的超额部分也不提供下一位驾驶者使用。
12. 构造一个有限状态机：当读取的输入符号所代表的数能够被 3 整除时，输出 1；否则输出 0。
13. 构造一个有限状态机：确定在输入串中当前所读取的最后一个符号是否为 1，且倒数第三个符号是否为 0。
14. 构造一个有限状态机：确定到目前为止所读取的输入串中，其结尾是否有至少 5 个连续的 1。
15. 构造一个有限状态机：确定到目前为止所读取的输入中，其最后的八个字符是否为computer。输入可能是任意的英文字母。

摩尔机 $M=(S, I, O, f, g, s_0)$ 由下列六部分构成：有限状态集 S；输入字母表 I；输出字母表 O；转移函数 f，它将每个由状态和输入组成的对映射为下一个状态；输出函数 g，它对每个状态指定一个输出；初始符号 s_0。摩尔机可以用状态表来表示，也可以用状态图来表示。状态表列出对应于每个状态和输入对的转移，以及对每个状态的输出。状态图画出状态、状态间的转移以及状态的输出。在状态图中，转移用带标号的箭头和输入表示，输出写在状态的旁边。

状态	f		
	输入		
	0	1	g
s_0	s_0	s_2	0
s_1	s_3	s_0	1
s_2	s_2	s_1	1
s_3	s_2	s_0	1

16. 构造具有右图中状态表的摩尔机的状态图。
17. 构造具有下列状态图的摩尔机的状态表。对每个输入串，摩尔机都产生一个输出串。特别对应于输入串 $a_1a_2\cdots a_k$ 的输出是 $g(s_0)g(s_1)\cdots g(s_k)$，其中 $s_i=f(s_{i-1}, a_i)$，$i=1, 2, \cdots, k$。



18. 对于下列每个输入串，求练习 16 中的摩尔机所生成的输出串。
    a)0101　　b)11111　　c)11101110111
19. 对于练习 18 中的每个输入串，求练习 17 中的摩尔机所生成的输出串。
20. 构造一个摩尔机：只要读取的输入符号的个数能够被 4 整除时，就输出 1。
21. 构造一个摩尔机，使其能够判断输入串是包含偶数个 1 还是奇数个 1，如果输入串中有偶数个 1，则输出 1；如果输入串中有奇数个 1，则输出 0。

11.3 不带输出的有限状态机

11.3.1 引言

有限状态机的最重要应用之一是语言识别。在设计和构造程序设计语言的编译器时，这个应用起着根本性的作用。在 11.2 节中，我们说明了可以用带输出的有限状态机来识别语言，方法是当读取的输入串在语言中时输出 1，否则输出 0。但是，还有一些其他类型的有限状态机，它们是为识别语言而专门设计的。这些机器不产生输出，但有终结状态。一个串能被它识别，当且仅当它把初始状态转变为终结状态之一。

11.3.2 串的集合

在讨论不带输出的有限状态机之前，先介绍一些关于串的集合的重要背景材料。这里定义的运算将广泛用于有限状态机识别的语言的讨论。

定义 1 设 V 是一个词汇表，A、B 是 V^* 的子集。A 和 B 的连接是所有形如 xy 的串构成的集合，记为 AB，其中 x 是 A 中的串，y 是 B 中的串。

例 1 设 $A=\{0, 11\}$，$B=\{1, 10, 110\}$。求 AB 和 BA。

解 集合 AB 包括所有 A 中串和 B 中串的连接，故 $AB=\{01, 010, 0110, 111, 1110,$

11110}。集合 BA 包括所有 B 中串和 A 中串的连接，故 $BA=\{10, 111, 100, 1011, 1100, 11011\}$。 ■

注意，对于字母表 V 和 V^* 的子集 A 与 B，$AB=BA$ 不一定成立，如例 1 所示。

由两个串集合连接的定义还可以定义 $A^n(n=0, 1, 2, \cdots)$。其递归定义如下：

$$A^0=\{\lambda\}$$

$$A^{n+1}=A^nA,\ n=0, 1, 2, \cdots$$

例 2 设 $A=\{1, 00\}$。当 $n=0, 1, 2, 3$ 时求 A^n。

解 易知：$A^0=\{\lambda\}$，$A^1=A^0A=\{\lambda\}A=\{1, 00\}$。为求 A^2，取 A 中元素对的连接。从而 $A^2=\{11, 100, 001, 0000\}$。为求 A^3，取 A^2 和 A 中的元素进行连接，由此得到 $A^3=\{111, 1100, 1001, 10000, 0011, 00100, 00001, 000000\}$。 ■

定义 2 设 A 是 V^* 的一个子集。A 的克莱因⊖闭包是 A 中任意多个串的连接组成的集合，记为 A^*，即 $A^*=\bigcup_{k=0}^{\infty}A^k$。

例 3 求集合 $A=\{0\}$，$B=\{0, 1\}$，$C=\{11\}$的克莱因闭包。

解 A 的克莱因闭包是 0 与自己的任意多次连接，故 $A^*=\{0^n \mid n=0, 1, 2, \cdots\}$。$B$ 的克莱因闭包是任意多个串的连接，但这些串只能是 0 或 1，因而这个闭包是字母表 $V=\{0, 1\}$上的所有串，即 $B^*=V^*$。最后，C 的克莱因闭包是 11 与自己的任意多次连接，所以，C^* 是由偶数个 1 组成的串的集合，即 $C^*=\{1^{2n} \mid n=0, 1, 2, \cdots\}$。 ■

11.3.3 有限状态自动机

现在给出不带输出的有限状态机的定义，这样的机器也叫*有限状态自动机*(finite-state automata)，这也是在本节中将使用的术语。(注：automata 的单数形式是 automaton。)这些机器与 11.2 节中研究的有限状态机不同，它们不产生输出，但它们有一个终结状态集合。我们将看到，它们识别的字符串是将初始状态变为终结状态的输入串。

定义 3 *有限状态自动机* $M=(S, I, f, s_0, F)$由下列五部分组成：*一个有限的状态集合* S，*一个有限的输入字母表* I，*一个转移函数* f(f 为每个状态和输入对指派下一个状态)，*一个初始状态* s_0，和*一个由终结状态构成的* S *的子集* F。

有限状态自动机也可用状态表或状态图来表示。在状态图中，终结状态用双圈表示。

例 4 构造有限状态自动机 $M=(S, I, f, s_0, F)$的状态图，其中 $S=\{s_0, s_1, s_2, s_3\}$，$I=\{0, 1\}$，$F=(s_0, s_3)$，转移函数 f 如表 11-6 所示。

解 所求的状态图如图 11-9 所示。注意：输入 0 和 1 都将 s_2 变为 s_0，所以从 s_2 到 s_0 的边上有 0 和 1。 ■



⊖ **克莱因(Stephen Cole Kleene，1909—1994)** 克莱因生于美国康涅狄格州的哈特福德。他的母亲艾丽丝·科尔(Alice Lena Cole)是一位诗人，他的父亲古斯塔夫·克莱因(Gustav Adolph Kleene)是一位经济学教授。克莱因曾就读于 Amherst 学院，1934 年在普林斯顿大学获博士学位，导师是著名逻辑学家丘奇(Alonzo Church)。1935 年，克莱因成为威斯康星大学的教员，除了有几次离开(包括去普林斯顿高等研究所)外，他一直待在那里。二战期间，他成为美国海军预备役军官学校的航海教师，后来成为海军研究实验室的主任。克莱因研究可计算性和可判定性问题，对递归函数论做出了重要贡献，并且证明了自动机理论中的一个中心结果。他曾是威斯康星大学的数学研究中心的代理主任和文理学院的院长。他还曾经学习过博物学，发现了一族以前没有描述过的蝴蝶，这族蝴蝶以他的名字命名。他还是一个热心的徒步旅行者和登山者。克莱因还因机敏地讲奇闻轶事而闻名，他的大嗓门在几间办公室之外都能听到。

表 11-6

状态	f	
	输入	
	0	1
s_0	s_0	s_1
s_1	s_0	s_2
s_2	s_0	s_0
s_3	s_2	s_1

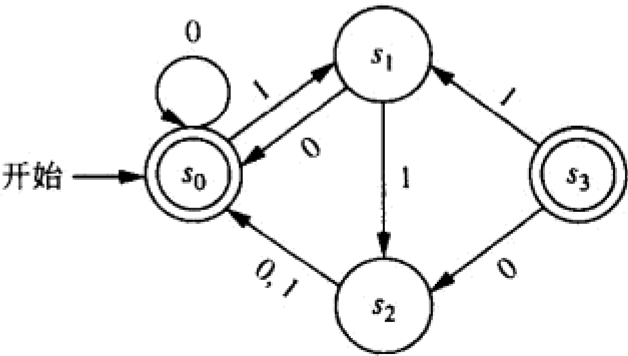


图 11-9 一个有限状态自动机的状态图

可以扩展转移函数，使其对所有状态与串的对都有值。设 $x=x_1x_2\cdots x_k$ 是 I^* 中的一个串，则 $f(s_1, x)$是这样得到的状态：以状态 s_1 开始，对 x 从左到右连续地使用其每个符号。从 s_1 我们用 $f(s_1, x)=f(s_k, x_k)$进入状态 $s_2=f(s_1, x_1)$，然后进入状态 $s_2=f(s_2, x_2)$，等等。

称串 x 可以被机器 $M=(S, I, f, s_0, F)$**识别**或**接受**，如果 x 将初始状态变为一个终结状态，即 $f(s_0, x)$是 F 中的一个状态。机器 M **识别**(或**接受**)的语言是 M 识别的所有串的集合，记为 $L(M)$。如果两个有限状态自动机识别相同的语言，则称它们是**等价的**。

例 5 求图 11-10 表示的有限状态自动机 M_1，M_2 和 M_3 所识别的语言。

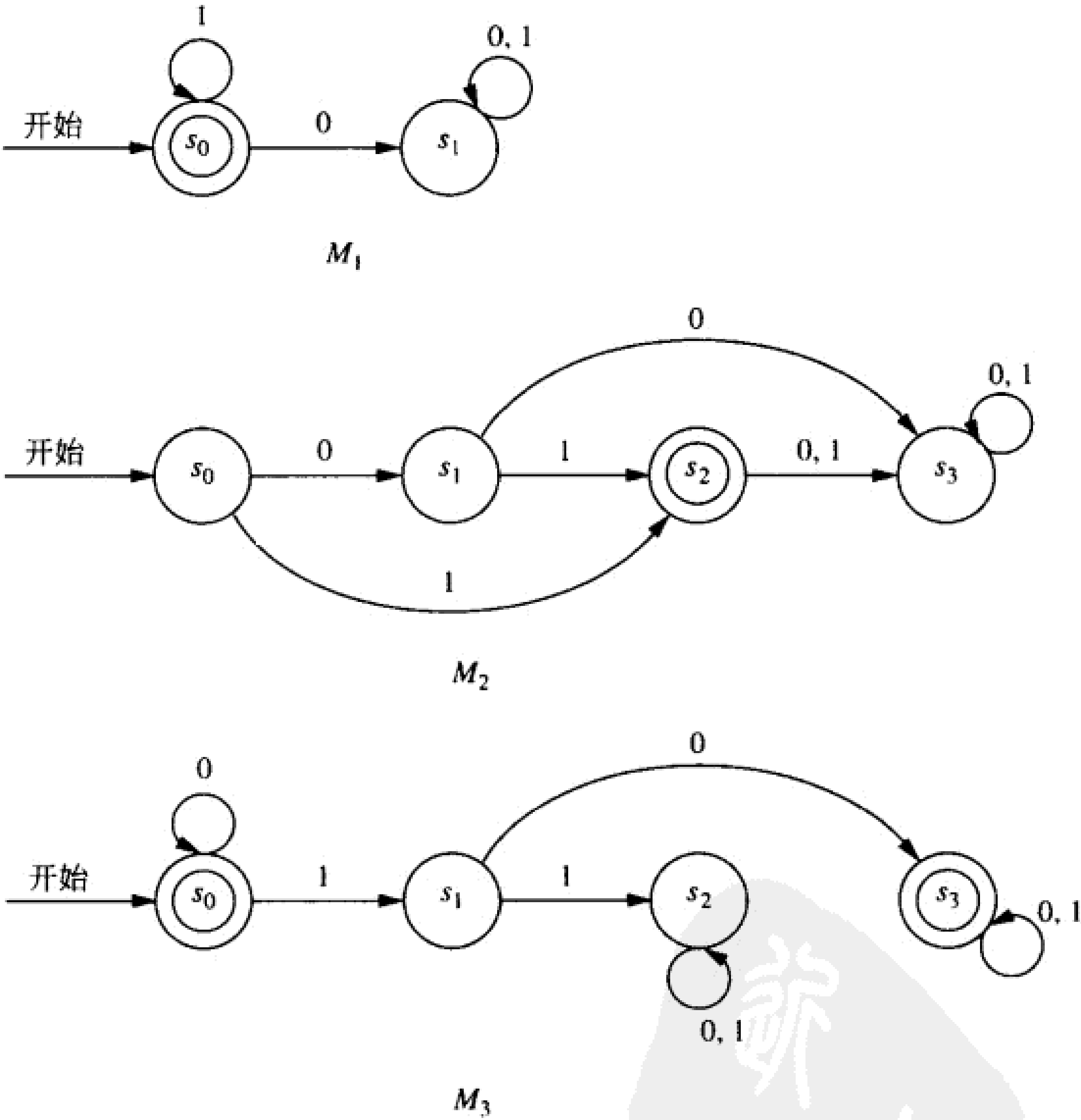


图 11-10 一些有限状态自动机

解 M_1 只有一个终结状态 s_0，而将 s_0 变为自身的串是由 0 个、1 个或更多连续 1 组成的串。所以 $L(M_1)=\{1^n \mid n=0, 1, 2, \cdots\}$。

M_2 只有一个终结状态 s_2，而将 s_0 变为 s_2 的串只有 1 和 01，所以 $L(M_2)=\{1, 01\}$。

M_3 的终结状态有 s_0 和 s_3，将 s_0 变为自身的串有 λ，0，00，000，…，即由零个以上(包括零个)连续的 0 构成的串。将 s_0 变为 s_3 的串只有这样的串：开头是零个以上(包括零个)连续的 0 构成的串，接着是 10，然后是任意的串。故 $L(M_3)=\{0^n, 0^n10x \mid n=0, 1, 2, \cdots, x$ 是任意的串$\}$。■

到目前为止所讨论的有限状态自动机都是确定型的，因为对每对状态和输入值，转移函数只给出唯一的下个一状态。还有一种重要的有限状态自动机，它对每对输入值和状态，有多个可能的下一个状态，这样的机器称为非确定型的。非确定型有限状态自动机在判断哪些语言可以由有限状态自动机识别中非常重要。

定义 4 非确定型有限状态自动机 $M=(S, I, f, s_0, F)$ 由下列五部分组成：一个状态的集合 S，一个输入字母表 I，一个转移函数 f(f 为每个状态和输入对指派一个状态集合)，一个初始状态 s_0，还有一个由终结状态构成的 S 的子集 F。

非确定型有限状态自动机也可用状态表和状态图来表示。在状态表中，对每对状态和输入值，列出所有可能的下一个状态。在状态图中，从一个状态到每个可能的下一个状态，都画一个边，这个边的标号是导致这个转移的一个或多个输入。

例 6 求状态表如表 11-7 所示的非确定型有限状态自动机的状态图。终结状态为 s_2 和 s_3。

解 这个自动机的状态图如图 11-11 所示。■

表 11-7

状态	f 输入 0	f 输入 1
s_0	s_0，s_1	s_3
s_1	s_0	s_1，s_3
s_2		s_0，s_2
s_3	s_0，s_1，s_2	s_1

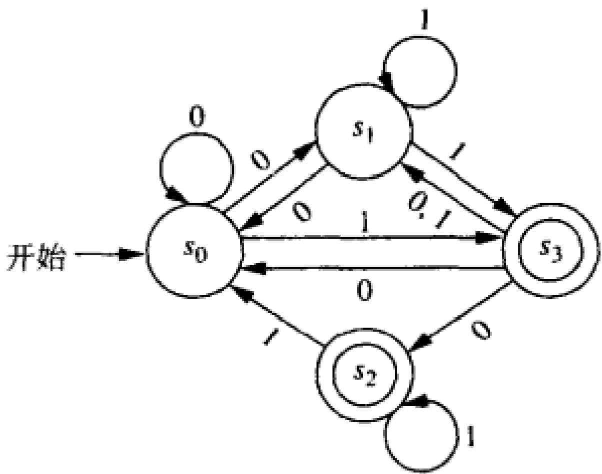


图 11-11 其状态表为表 11-7 的非确定型有限状态自动机

例 7 求状态图如图 11-12 所示的非确定型有限状态自动机的状态表。

解 这个自动机的状态表如表 11-8 所示。■

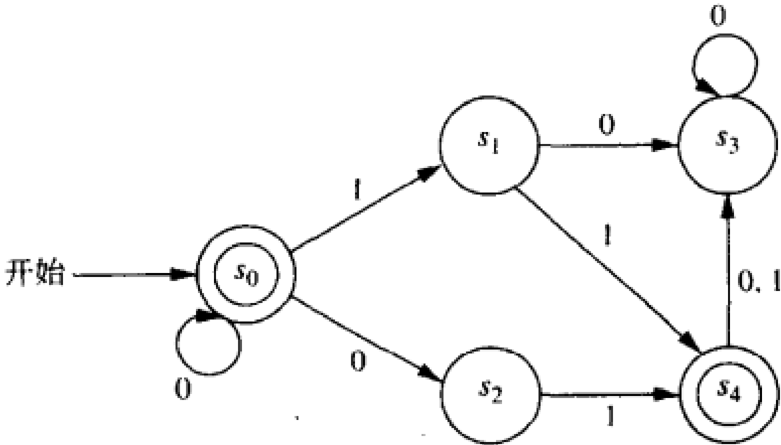


图 11-12 一个非确定型有限状态自动机

表 11-8

状态	f 输入 0	f 输入 1
s_0	s_0，s_2	s_1
s_1	s_3	s_4
s_2		s_4
s_3	s_3	
s_4	s_3	s_3

非确定型有限状态自动机怎么识别串 $x=x_1x_2\cdots x_k$ 呢？第一个输入符号 x_1 将初始状态 s_0 变为状态集合 s_1。下一个输入符号 x_2 将 s_1 中的每个状态都变为一个状态集合。设 s_2 是这些集合的并集。将这个过程继续下去，在一个步骤中，对上一个步骤产生的每个状态和当前的输入符号，都要求其产生的状态。使用 x 从 s_0 所能得到的状态中，如果有一个终结状态，我们就识别或接受串 x。非确定型有限状态自动机所识别的语言是这个自动机所识别的所有串的集合。

例 8 求图 11-12 所示的非确定型有限状态自动机所识别的语言。

解 因为 s_0 是终结状态，且当输入是 0 时，有从 s_0 到自身的转移，故此机器识别所有零个或更多个连续的 0 组成的串。因为 s_4 也是终结状态，对于任何串，若以此串作为输入时从 s_0 所能达到的状态集中包含 s_4，则此串就能被识别。这样的串只有：零个或更多个连续的 0 和后面跟 01 或 11 组成的串。因为 s_0 和 s_4 是仅有的终结状态，所以此机器识别的语言为 $\{0^n, 0^n01, 0^n11 \mid n\geqslant 0\}$。■

一个重要的事实是：非确定型有限状态自动机所能识别的语言也能被确定型有限状态自动机所识别。在下节中，我们将利用这个事实，以确定有限状态自动机能够识别哪些语言。

定理 1 如果语言 L 可以由非确定型有限状态自动机 M_0 所识别，则 L 也可以由一个确定型有限状态自动机 M_1 来识别。

证 我们来描述怎么从 M_0 构造识别 L 的确定型有限状态自动机 M_1。M_1 的每个状态都由 M_0 的状态集的一个子集构成，M_1 的初始符号是 $\{s_0\}$，即 M_0 的初始符号是 s_0 构成的集合。M_1 的输入集合与 M_0 的相同。对于 M_1 的一个给定状态 $\{s_{i_1}, s_{i_2}, \cdots, s_{i_k}\}$，输入串 x 将这个状态变为这个集合中元素的下一个状态构成的集合的并，即集合 $f(s_{i_1}, x)$，$f(s_{i_2}, x)$，$\cdots$，$f(s_{i_k}, x)$ 的并。M_1 的状态是 S 的所有子集，这里 S 是 M_0 的状态集。(如果此非确定型机器有 n 个状态，则确定型机器就有 2^n 个状态，因为所有子集都可以作为状态，包括空集，尽管实际使用的状态却很少。)如果 M_0 的一个子状态集含有终结状态，则它就为 M_1 的终结状态。

假设一个输入串可以由 M_0 识别，则用这个串从 s_0 出发可以达到的状态中有一个终结状态(读者应能对这一点作归纳证明)。这意味着：在 M_1 中，这个串能将 $\{s_0\}$ 引导至这样一个状态集：它是 M_0 的状态集的一个子集，且包含一个终结状态。这个子集是 M_1 的一个终结状态，所以这个串也能由 M_1 识别。还有，如果一个输入串不能由 M_0 识别，则它也就不能导致 M_0 中的任何终结状态。(读者应该能够给出它的详细证明。)因而这个输入串也不可能由 $\{s_0\}$ 导致 M_1 中的一个终结状态。□

例 9 求一个确定型有限状态自动机，能与例 7 中的非确定型有限状态自动机识别相同的语言。

解 所求的确定型有限状态自动机如图 11-13 所示，它是根据例 7 中的非确定型有限状态自动机构造的。此确定型自动机的状态是那个非确定型机器的状态集的子集。对于一个输入符号，

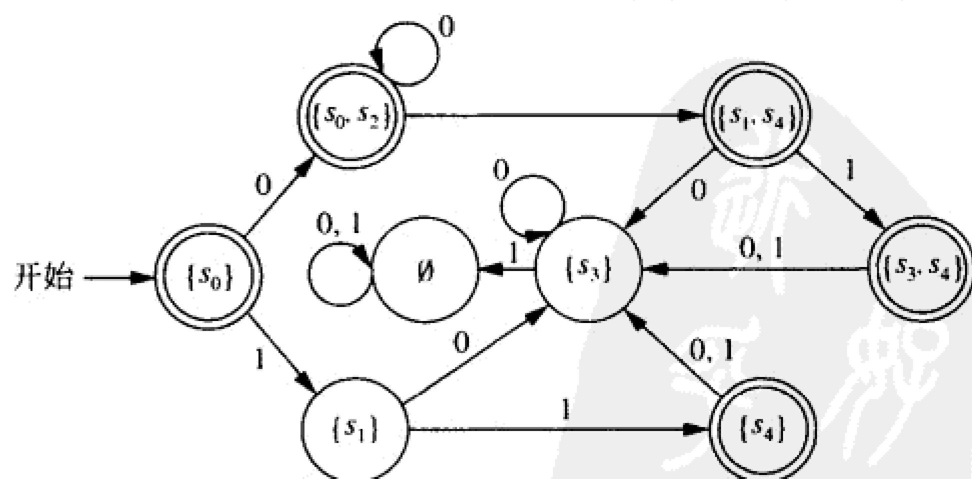


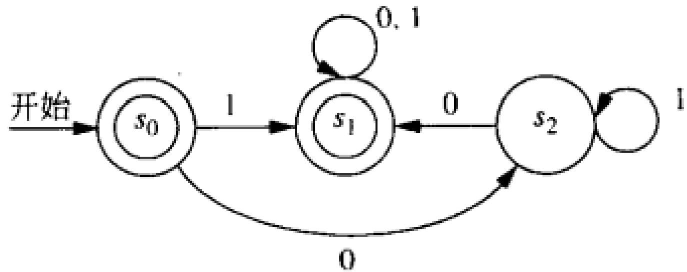
图 11-13 与例 7 中非确定型有限状态自动机等价的确定型有限状态自动机

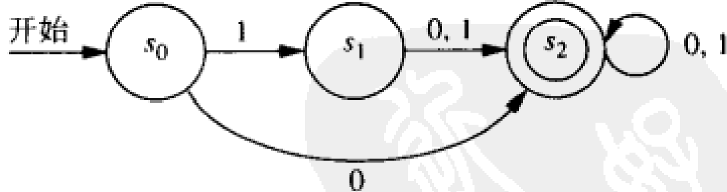
一个子集的下一个状态是这样的集合：它由这个子集中的元素在那个非确定型机器中的所有下一个状态所构成。例如，对于输入 0，$\{s_0\}$转为$\{s_0, s_2\}$，因为在那个非确定型机器中，s_0 有到它自己和 s_2 的转移；集合$\{s_0, s_2\}$对输入 1 转为$\{s_1, s_4\}$，因为在那个非确定型机器中，对输入 1 s_0 只转为 s_1，s_2 只转为 s_4；集合$\{s_1, s_4\}$对输入 0 转为$\{s_3\}$，因为在那个非确定型机器中，对输入 0 s_1 和 s_4 都只转为 s_3。所有以这种方法得到的子集都包括在这个确定型有限状态机器里。注意：空集也是这个机器的一个状态，因为它是$\{s_3\}$对输入 1 的所有下一个状态构成的子集。初始状态是$\{s_0\}$，终结状态是那些包含 s_0 或 s_4 的集合。■

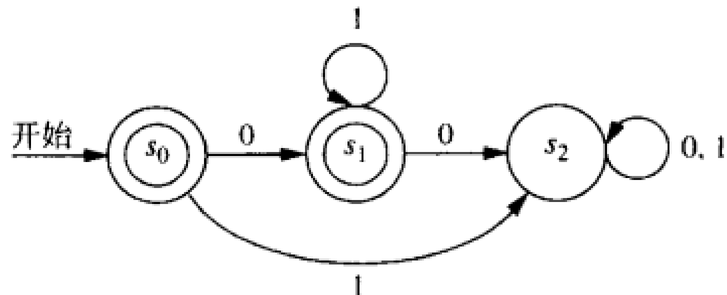
练习

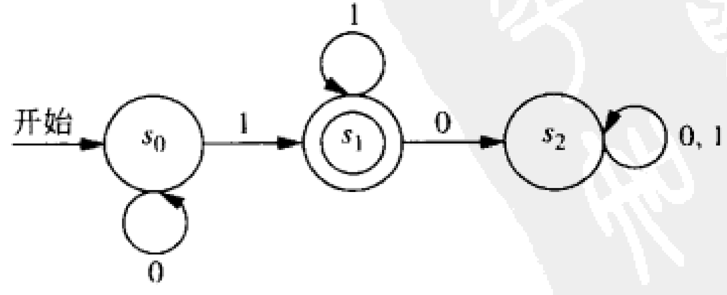
1. 设 $A=\{0, 11\}$，$B=\{00, 01\}$，求下列集合：
   a) AB　b) BA　c) A^2　d) B^3
2. 设 A 是串的一个集合。证明 $A\emptyset=\emptyset A=\emptyset$。
3. 求所有串的集合对 A 和 B 使得 $AB=\{10, 111, 1010, 1000, 10111, 101000\}$。
4. 证明下列等式成立：
   a) $\{\lambda\}^*=\{\lambda\}$　b) 对任意串的集合 A，$(A^*)^*=A$
5. 对于下列集合 A，描述集合 A^* 的元素：
   a) $\{10\}$　b) $\{111\}$　c) $\{0, 01\}$　d) $\{1, 101\}$
6. 设 V 是一个字母表，A 和 B 是 V^* 的子集。证明：$|AB|\leqslant A|\ |B|$。
7. 设 V 是一个字母表，A 和 B 是 V^* 的子集，且 $A\subseteq B$。证明：$A^*\subseteq B^*$。
8. 设 V 是一个字母表，A 是 V^* 的子集。证明下列命题是正确的或是错误的：
   a) $A\subseteq A^2$　b) 如果 $A=A^2$，则 $\lambda\in A$　c) $A\{\lambda\}=A$
   d) $(A^*)^*=A^*$　e) $A^*A=A^*$　f) $|A^n|=|A|^n$
9. 确定下列集合是否包含串 11101。
   a) $\{0, 1\}^*$　b) $\{1\}^*\{0\}^*\{1\}^*$　c) $\{11\}\{1\}^*\{01\}$
   d) $\{11\}^*\{01\}^*$　e) $\{111\}^*\{0\}^*\{1\}$　f) $\{111, 000\}\{00, 01\}$
10. 确定下列串能否由图 11-9 中的确定型有限状态自动机所识别。
    a) 010　b) 1101　c) 1111110　d) 010101010
11. 对于下列每个集合，确定其中的每个串是否都能由图 11-9 中的确定型有限状态自动机所识别。
    a) $\{0\}^*$　b) $\{0\}\{0\}^*$　c) $\{1\}\{0\}^*$
    d) $\{01\}^*$　e) $\{0\}^*\{1\}^*$　f) $\{1\}\{0, 1\}^*$

在练习 12～16 中，求所给的确定型有限状态自动机所识别的语言。

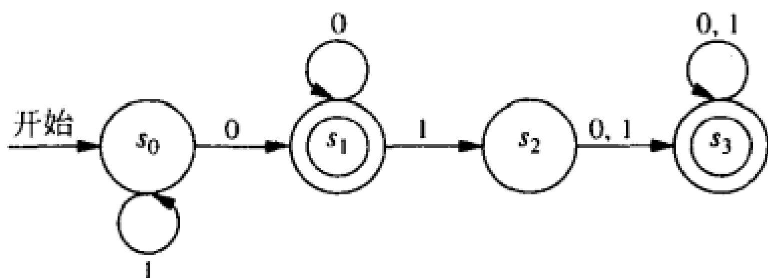
12.


13.


14.


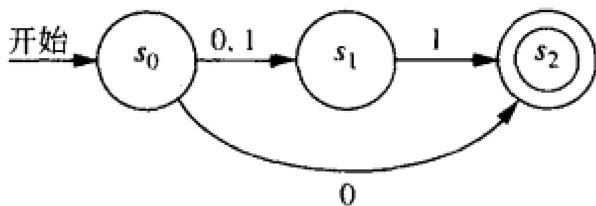
15.


16.

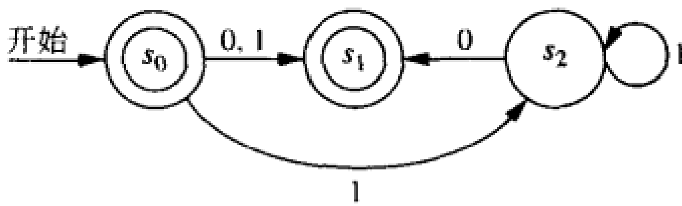


在练习 17～21 中，求所给的非确定型有限状态自动机所识别的语言。

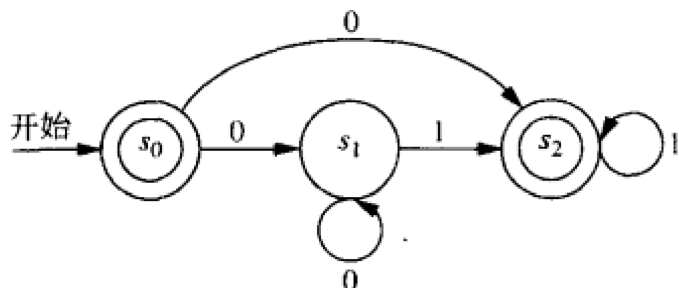
17.



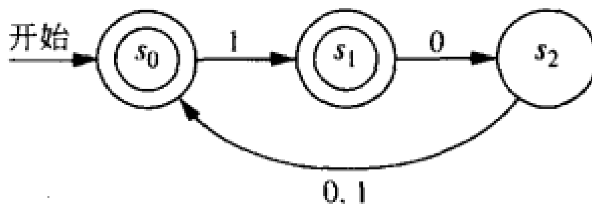
18.



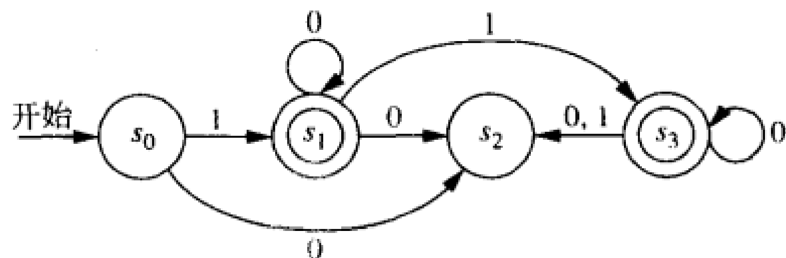
19.



20.



21.



22. 求一个确定型有限状态自动机，能与练习 17 中的非确定型有限状态自动机识别相同的语言。
23. 求一个确定型有限状态自动机，能与练习 18 中的非确定型有限状态自动机识别相同的语言。
24. 求一个确定型有限状态自动机，能与练习 19 中的非确定型有限状态自动机识别相同的语言。
25. 求一个确定型有限状态自动机，能与练习 20 中的非确定型有限状态自动机识别相同的语言。
26. 求一个确定型有限状态自动机，能与练习 21 中的非确定型有限状态自动机识别相同的语言。
27. 对于下列每个集合，分别求一个识别它的确定型有限状态自动机：

a) $\{0\}$ b) $\{1, 00\}$ c) $\{1^n \mid n=2, 3, 4, \cdots\}$

28. 对于练习 27 中的每种语言，分别求一个识别它的非确定型有限状态自动机，并且，如果可能的话，使之所具有的状态比你在练习 27 中所给的确定型有限状态自动机更少。

*29. 对于由个数相同的 0 和 1 组成的串构成的集合，证明没有有限状态自动机能够识别它。

11.4 语言的识别

11.4.1 引言

我们已经知道，有限状态自动机可以用作语言识别器。那么这些机器能识别什么样的集合呢？这个问题虽然看起来极为困难，但有限状态自动机所识别的语言还是有一个简单特征。这个问题由美国数学家克莱因(Stephen Kleene)于 1956 年首先解决。他证明了：一个集合能够被一个有限状态自动机识别当且仅当这个集合是这样得到的：以任意顺序通过对空集、空串和单字符串的连接、并或克莱因闭包构造出来。以这种方法构造出来的集合称为正则集合。

在 11.1 节中我们定义了正则文法。从名称上可以想到，有限状态自动机所识别的正则集合与正则文法间具有某种联系。特别是，一个集合是正则的当且仅当它可以由一个正则文法生成。

但也有一些集合不能由任何有限状态自动机识别，我们将给出这样一个集合。在本节的最后我们将简短讨论一些更强大的计算模型，如下推自动机和图灵机。

11.4.2 正则集合

正则集合是这样集合：从空集、空串、单字符串开始，以任意顺序通过连接、并和克莱因闭包运算形成的。我们将知道，正则集合正是有限状态自动机识别的语言的集合。为定义正则集合，首先要定义正则表达式。

定义 1 集合 I 上的正则表达式的递归定义如下：

符号 $\emptyset$ 是一个正则表达式；

符号 λ 是一个正则表达式；

若 $x \in I$ 时，符号 x 是一个正则表达式；

当 **A**，**B** 是正则表达式时，符号(**AB**)，(**A**$\cup$**B** 和 **A*** 都是正则表达式。

每个正则表达式都表示一个由下列规则指出的集合：

$\emptyset$ 表示空集，即没有串的集合；

λ 表示集合$\{\lambda\}$，即空串组成的集合；

x 表示集合$\{x\}$，它只包含单个符号 x 组成的串；

(**AB**)表示 **A** 和 **B** 代表的集合的连接；

(**A**$\cup$**B**)表示 **A** 和 **B** 代表的集合的并；

A* 表示 **A** 代表的集合的克莱因闭包。

正则表达式表示的集合称为*正则集合*。今后正则集合将由正则表达式来描述，所以，当我们提到正则集合 **A** 时，指的是此正则表达式 **A** 表示的集合。下面的例子说明了怎么用正则表达式来规定正则集合。

例 1 正则表达式 **10***，(**10**)*，**0**$\cup$**01**，**0**(**0**$\cup$**1**)*和(**0*****1**)*所规定的正则集合中有哪些串？

解 这些正则表达式所表示的正则集合如表 11-9 所示。读者可以自己验证。■

表 11-9

表达式	串	表达式	串
10*	1 后面跟任意多个 0(也可以没有 0)	**0**(**0**$\cup$**1**)*	以 0 开头的任意串
(**10**)*	10 的任意多次复制(包括空串)	(**0*****1**)*	不以 0 结尾的任意串
0$\cup$**01**	串 0 或 01		

11.4.3 克莱因定理

在 1956 年，克莱因证明了正则集合就是有限状态自动机识别的集合。因此，这个重要结论称为克莱因定理。

定理 1 *克莱因定理* 一个集合是正则的，当且仅当它可由一个有限状态自动机识别。

克莱因定理是自动机理论的中心定理之一，我们只证明必要性部分，即证明每个正则集合都可由一个有限状态自动机识别。充分性部分(有限状态自动机识别的集合都是正则的)留作练习。

证 回忆一下正则集合是通过正则表达式定义的，而正则表达式是递归定义的。所以，如果我们能证明下列事情，就证明了每个正则集合都可由一个有限状态自动机识别。

1. 证明 $\emptyset$ 可由一个有限状态自动机识别。

2. 证明{λ}可由一个有限状态自动机识别。

3. 证明{a}可由一个有限状态自动机识别，其中 a 是 I 中的符号。

4. 当 A 和 B 都可由有限状态自动机识别时，证明 AB 也可由有限状态自动机识别。

5. 当 A 和 B 都可由有限状态自动机识别时，证明 $A\cup B$ 也可由有限状态自动机识别。

6. 当 A 可由有限状态自动机识别时，证明 A^* 也可由有限状态自动机识别。

下面分别讨论每个任务。首先证明 $\varnothing$ 可以由有限状态自动机来识别。为此，需要构造一个没有终结状态的自动机。图 11-14a 就是这样一个自动机。

其次证明{λ}也可由有限状态自动机来识别。为此需要构造一个识别空串 λ 的自动机，且除 λ 外，它不识别任何其他的串。这个自动机可以这样来构造：初始状态 s_0 也用作终结状态，且对任何其他的串，没有转移能将 s_0 变为一个终结状态。图 11-14b 就是这样一个非确定型的自动机。

第三，证明{a}也可由非确定型有限状态自动机来识别。为此构造如下机器：初始状态是 s_0，终结状态是 s_1，对于输入 a 有一个从 s_0 到 s_1 的转移，且没有其他转移。这个机器识别的唯一的串是 a。这个机器如图 11-14c 所示。

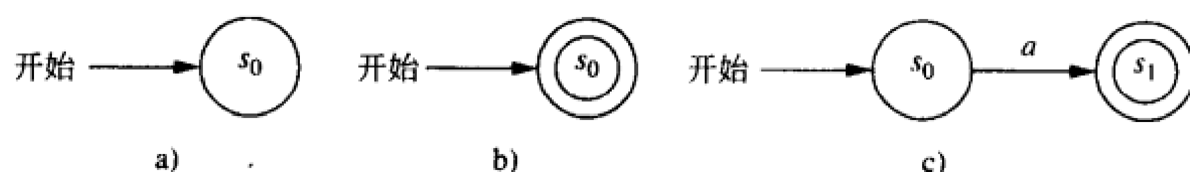


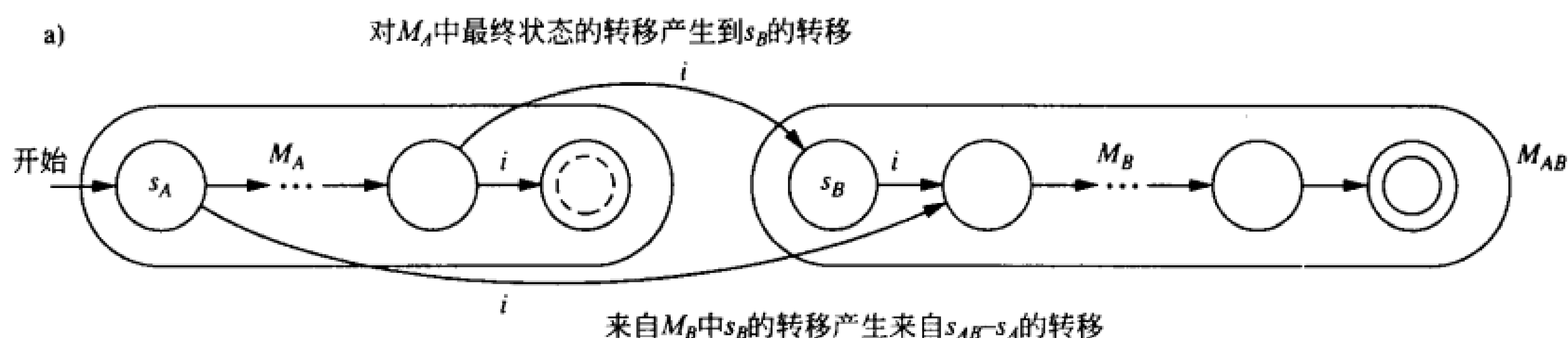
图 11-14　识别某些基本集合的非确定型有限状态自动机

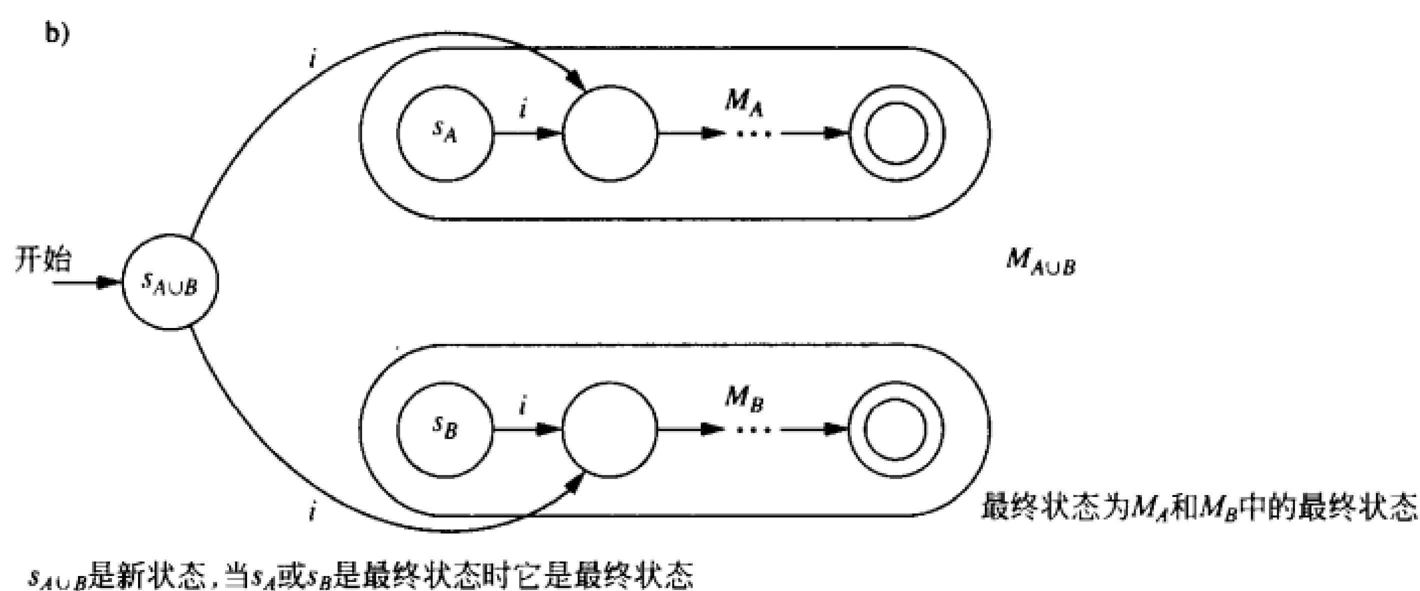
下面证明：如果 A 和 B 都是可由有限状态自动机识别的语言，则 AB 和 $A\cup B$ 也可由有限状态自动机识别。设 A 是由 $M_A=(S_A, I, f_A, s_A, F_A)$识别的，$B$ 是由 $M_B=(S_B, I, f_B, s_B, F_B)$识别的。

先来构造识别 A 与 B 的连接 AB 的有限状态自动机 $M_{AB}=(S_{AB}, I, f_{AB}, s_{AB}, F_{AB})$，它是由识别 A 和 B 的两个机器串联而成的，使得 A 中的串将这个组合机器从 M_A 的初始状态 s_A 变为 M_B 的初始状态 s_B。B 中的串应该将这个组合机器从 s_B 变为此组合机器的一个终结状态。因此我们进行如下构造：令 S_{AB} 是 $S_A\cup S_B$；初始状态 s_{AB} 就是 s_A；终结状态集 F_{AB} 或者是 M_B 的终结状态集，或者当且仅当 $\lambda\in A\cap B$ 时还包括 s_{AB}；M_{AB}的转移除了包括 M_A 和 M_B 中的全部转移之外，还有一些新的转移。对于 M_A 中每个导致终结状态的转移，在 M_{AB} 中增加一个在同一个输入上从同一个状态到 s_B 的转移，这样，A 中的串就将 M_{AB} 从 s_{AB} 变为 s_B，然后 B 中的串将 s_B 变为 M_{AB} 的一个终结状态。而且，对每个从 s_B 出发的转移，还在 M_{AB} 中增加一个从 s_{AB} 到同一个状态的转移。图 11-15a 包括了这个构造的所有说明。

现在构造识别 $A\cup B$ 的机器 $M_{A\cup B}=\{S_{A\cup B}, I, f_{A\cup B}, s_{A\cup B}, F_{A\cup B}\}$。这个自动机是将 M_A 和 M_B 并行组合起来的，它使用一个新的初始状态，此初始状态具有 s_A 和 s_B 所具有的转移。令 $S_{A\cup B}=S_A\cup S_B\cup\{s_{A\cup B}\}$，其中 $s_{A\cup B}$是$M_{A\cup B}$的新初始状态。当 $\lambda\in A\cup B$ 时，令终结状态集 $F_{A\cup B}$是$F_A\cup F_B\cup\{s_{A\cup B}\}$，否则为 $F_A\cup F_B$。$M_{A\cup B}$的转移除了包括M_A 和M_B 中的所有转移外，还有如下的转移：对输入 i 从 s_A 到状态 s 的每个转移，包括一个对输入 i 从 $s_{A\cup B}$到状态 s 的转移；对输入 i 从 s_B 到状态 s 的每个转移，包括一个在输入 i 从 $s_{A\cup B}$到状态 s 的转移。这样，在这个新机器中，A 中的串将从 $s_{A\cup B}$ 导致一个终结状态，B 中的串也将从 $s_{A\cup B}$ 导致一个终结状态。图 11-15b 说明了 $M_{A\cup B}$的构造。

最后构造识别 A^* 的机器 $M_{A^*}=(S_{A^*}, I, f_{A^*}, s_{A^*}, F_{A^*})$，其中 A^* 是 A 的克莱因闭包。





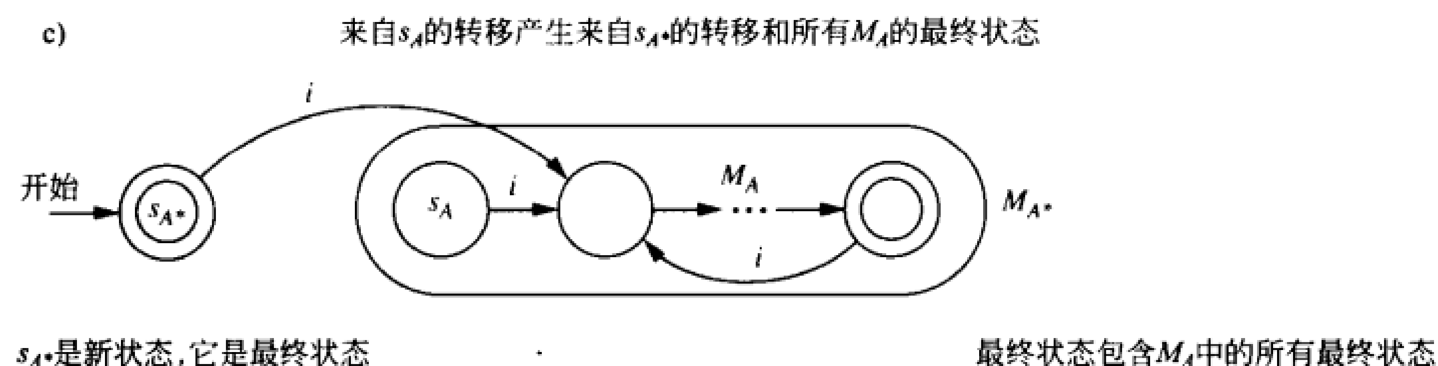


图 11-15　构造识别连接、并以及克莱因闭包的自动机

令 S_{A^*} 包含 S_A 中所有状态，还有一个状态 s_{A^*}，它是这个新机器的初始状态。终结状态集 F_{A^*} 包含了 F_A 中所有状态和初始状态 s_{A^*}，因为必须要识别 λ。为了识别 A 中任意多个串的连接，我们包括 M_A 中的所有转移，以及如下转移：与从 s_A 出发的转移相匹配的从 s_{A^*} 出发的转移[⊖]，与从 s_A 出发的转移相匹配的从每个终结状态出发的转移。有了这个转移集，对于由 A 中的一些串连接而成的串，当 A 中的第一个串读完时，它将 s_A^* 变到一个终结状态，当 A 中的第二个串读完时，它又回到一个终结状态，等等。图 11-15c 说明了我们刚才的构造。□

用以上描述的过程，可以对任意正则集合构造一个非确定型有限状态自动机。下面用例子来说明这一点。

例 2　构造一个非确定型有限状态自动机来识别正则集合 $\mathbf{1}^* \cup \mathbf{01}$。

解　首先构造一个机器来识别 $\mathbf{1}^*$。为此，先构造一个识别 $\mathbf{1}$ 的机器，再使用在定理 1 的证

⊖ 即它们是在相同输入上到相同状态的转移。——译者注

明中描述的构造 M_{A^*} 的方法。第二步，构造识别 **01** 的机器。先分别构造识别 **0** 和 **1** 的机器，再使用在定理 1 的证明中描述的构造 M_{AB} 的方法。最后，用在定理 1 的证明中描述的构造 $M_{A\cup B}$ 的文法，构造识别 **1* ∪01** 的机器。所构造的有限状态自动机如图 11-16 所示。上述各个机器中的状态被标以不同的下标，即使对从前机器中继承下来的状态也是如此。注意：这样构造的机器并不是识别合 **1* ∪01** 的最简单的机器。图 11-16b 是识别同一个集合但简单得多的机器。■

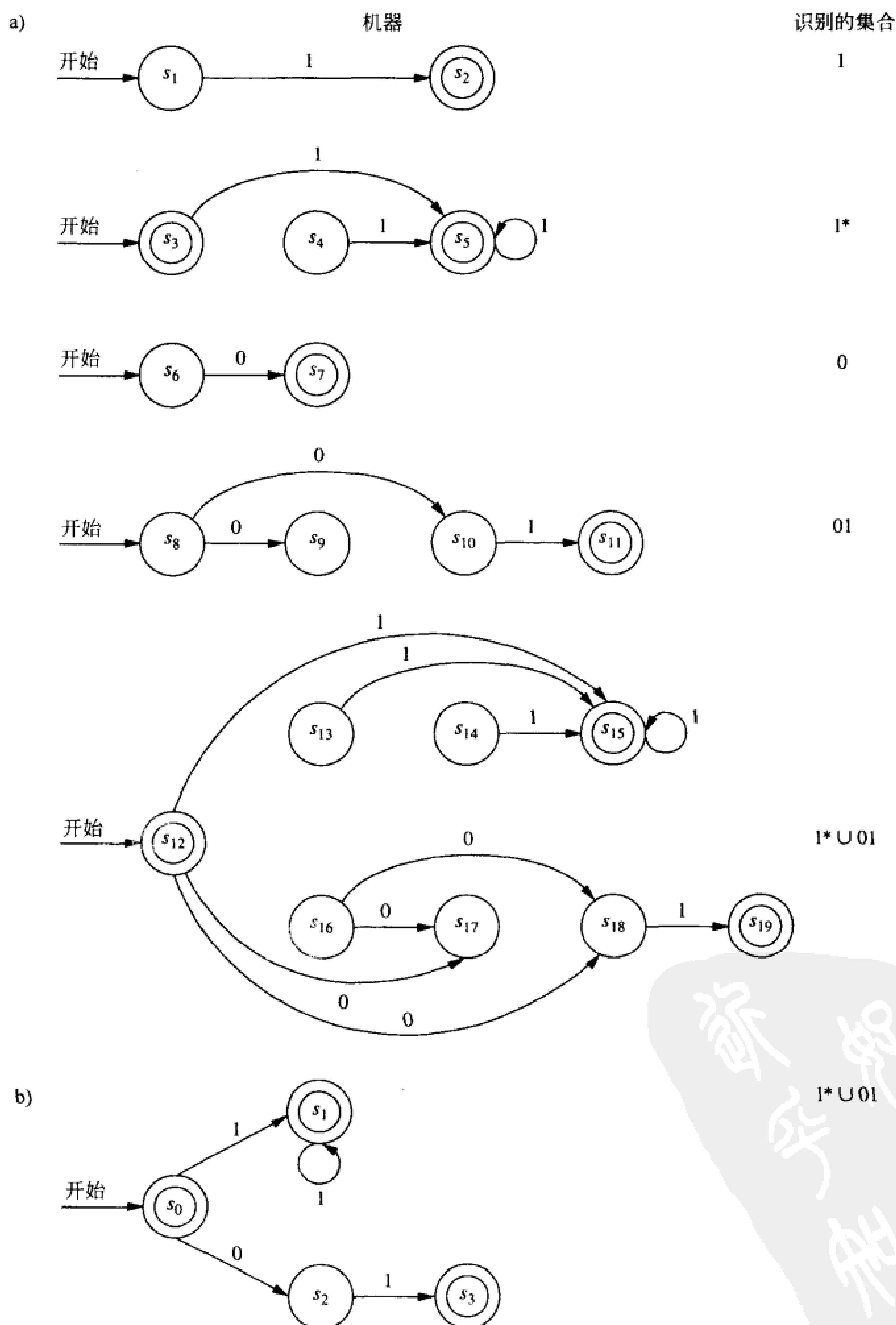


图 11-16　识别 **1* ∪01** 的非确定型有限状态自动机

11.4.4 正则集合和正则文法

在 11.1 节中介绍了短语结构文法，还定义了各种不同类型的文法。特别定义了正则文法(或 3 型文法)，这是一个形如 $G=(V, T, S, P)$ 的文法，文法的每个产生式的形式是 $S\to\lambda$，$A\to a$ 或 $A\to aB$，其中 a 是终结符，A 和 B 是非终结符。正如名称所暗示，正则文法和正则集合之间具有紧密的联系。

定理 2 一个集合可以由正则文法生成，当且仅当它是一个正则集合。

证 首先证明正则文法生成的集合是一个正则集合。设 $G=(V, T, S, P)$ 是一个正则文法，其生成的集合是 $L(G)$。为证明 $L(G)$ 是正则的，我们只要构造一个非确定型有限状态自动机 $M=(S, I, f, s_0, F)$ 来识别 $L(G)$。对 G 的每个非终结符 A，状态集 S 都包含一个相应的状态，S 还包含一个状态 s_F，它是一个终结状态。初始状态 s_0 是从初始符号 S 构造的。M 的转移是根据 G 的产生式按以下方式构造的：如果 $A\to a$ 是一个产生式，则包括一个对输入 a 从 s_A 到 s_F 的转移；如果 $A\to aB$ 是一个产生式，则包括一个在输入 a 上从 s_A 到 s_B 的转移。终结状态集包括 s_F，如果 $S\to\lambda$ 是 G 中产生式，则还要包括 s_0。不难证明，M 识别的语言与文法 G 生成的语言相等，即 $L(M)=L(G)$。这只要确定导致终结状态的词即可。详细证明留作练习。 □

在给出反方向的证明之前，先说明怎么构造一个非确定型机器，能识别由一个正则文法识别的集合。

例 3 构造一个非确定型有限状态自动机，使之识别正则文法 $G=(V, T, S, P)$ 生成的语言，其中 $V=\{0, 1, A, S\}$，$T=\{0, 1\}$，P 中的产生式为 $S\to 1A$，$S\to 0$，$S\to\lambda$，$A\to 0A$，$A\to 1A$ 和$A\to 1$。

解 图 11-17 是识别 $L(G)$的非确定型有限状态自动机的状态图。这个自动机是根据上面证明描述的过程构造的。在这个自动机中，s_0 是对应 S 的状态，s_1 是对应 A 的状态，s_2 是终结状态。 ■

现在来完成定理 2 的证明。

证 现在证明，如果一个集合是正则的，则存在一个正则文法生成它。设 M 是识别这个集合的一个有限状态机器，且具有性质：M 的初始状态 s_0 对任何转移都不是下一个状态。(可以根据练习 14 找到这样机器。)文法 $G=(V, T, S, P)$的定义如下：G 之符号集 V 是这样形成的，对 S 的每个状态和 I 中的每个输入符号，指派 V 中一个符号。G 的终结符集 T 是 G 中这样符号构成的，它是根据 I 中输入符号构造的符号。初始符号 S 是根据初始状态 s_0 构造的符号。G 的产生式集 P 是根据 M 中的转移构造的。特别地，如果状态 s 对输入 a 变到一个终结状态，则 P 中就包括产生式 $A_s\to a$，其中 A_s 是根据状态 s 构造的非终结符。如果状态 s 对输入 a 变到状态 t，则 P 中就包括产生式 $A_s\to aA_t$。P 中包括产生式 $S\to\lambda$，当且仅当 $\lambda\in L(M)$。因为 G 的产生式对应于 M 的转移，且导致终结符的产生式对应于导致终结状态的转移，因而不难证明 $L(G)=L(M)$。详细证明留作练习。 □

下面的例子说明怎么根据自动机来构造文法，使得该文法生成的语言就是这个自动机识别的语言。

例 4 求一个正则文法，使之生成的集合是图 11-18 表示的有限状态自动机所识别的正则集合。

解 生成该自动机所识别的集合的文法为 $G=(V, T, S, P)$，其中 $V=\{S, A, B, 0, 1\}$，其符号 S，A，B 分别对应于状态 s_0，s_1 和 s_2，$T=\{0, 1\}$，S 是初始符号，产生式为 $S\to 0A$，$S\to 1B$，$S\to 1$，$S\to\lambda$，$A\to 0A$，$A\to 1B$，$A\to 1$，$B\to 0A$，$B\to 1B$ 和 $B\to 1$。 ■

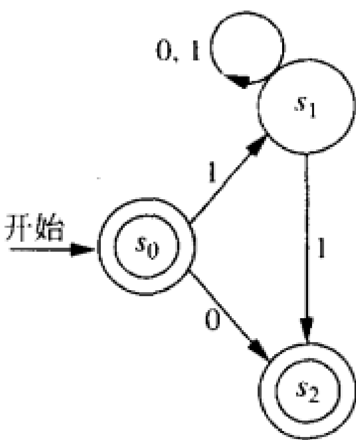


图 11-17 识别 $L(G)$的非确定型有限状态自动机

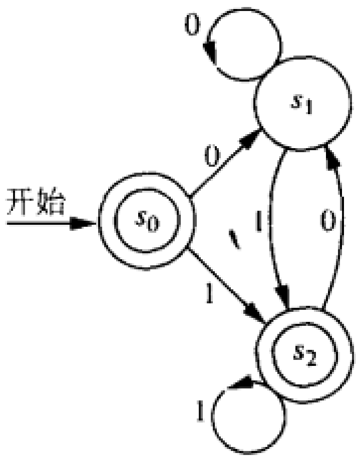


图 11-18 一个有限状态自动机

11.4.5 一个不能由有限状态自动机识别的集合

我们知道，一个集合能够由有限状态自动机识别当且仅当它是正则的。现在来证明存在不是正则的集合，方法是给出这样一个集合。证明这个集合不是正则的技术是一个重要方法，可以用来证明某类集合是不正则的。

例 5 集合$\{0^n1^n \mid n=0, 1, 2, \cdots\}$是所有如下的串构成的：先是一列 0，后跟同样个数的 1。证明这个集合不是正则的。

解 假如这个集合是正则的，则存在一个有限状态自动机 $M=(S, I, f, s_0, F)$识别它。设 N 是这个机器中状态的个数，即 $N=|S|$。因为 M 能识别所有这样构成的串：先是一列 0，后跟同样个数的 1，故它必定能识别 0^N1^N。设 s_0，s_1，s_2，…，s_{2N}是如下得到的状态序列：以 s_0 开始，以 0^N1^N 中的符号作为输入，且满足 $s_1=f(s_0, 0)$，$s_2=f(s_1, 0)$，…，$s_N=f(s_{N-1}, 0)$，$s_{N+1}=f(s_N, 1)$，…，$s_{2N}=f(s_{2N-1}, 1)$。注意 s_{2N}是一个终结状态。

因为只有 N 个状态，根据鸽巢原理，在 s_0，…，s_N 这头 $N+1$ 个状态中，至少有两个是相同的。假设 s_i，s_j 是两个这样相同的状态$(0\leqslant i<j\leqslant N)$，则这表示 $f(s_i, 0^t)=s_j$，其中 $t=i-j$。由此可知，t 次使用输入 0 后，可以得到一个从 s_i 回到它自己的循环，如状态图 11-19 所示。

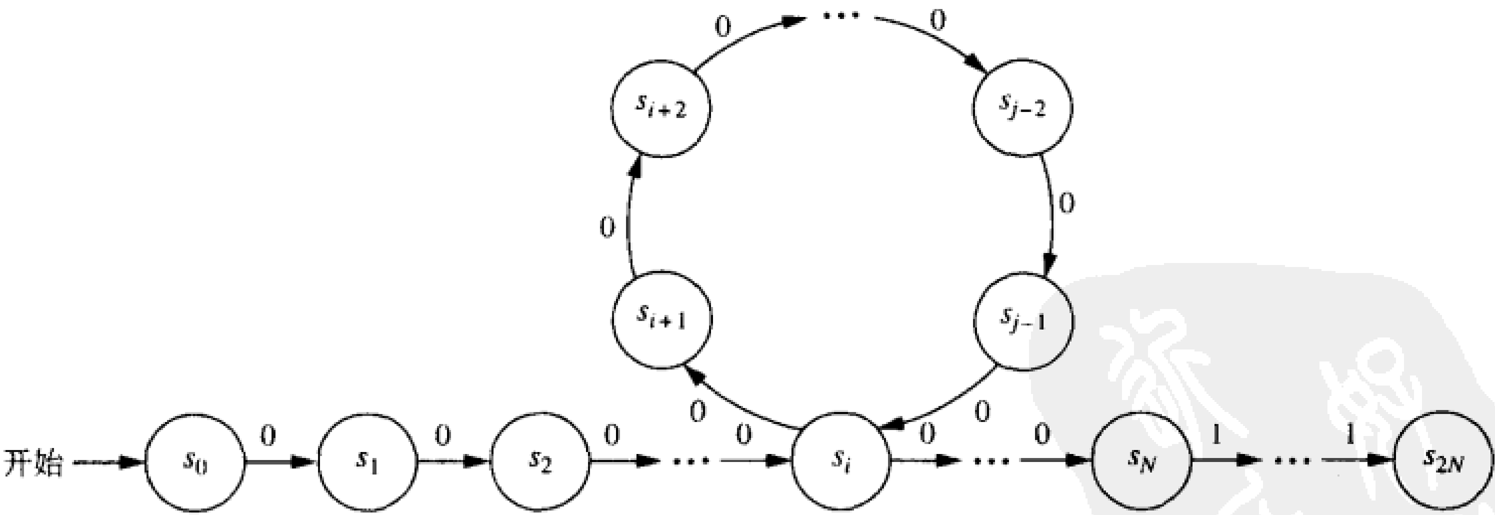


图 11-19 0^N1^N 所产生的路径

现在考虑输入串 $0^N0^t1^N=0^{N+t}1^N$。此串的前半部分比后半部分多了 t 个连续的 0。因为这个串不具有形式 0^n1^n(因为其中 0 的个数比 1 的个数多)，故不能被 M 识别，因而$f(s_0, 0^{N+t}1^N)$也就不是终结状态。但当用 $0^{N+t}1^N$ 作为输入时，得到的结束状态与以前一样，即 s_{2N}。其理由是：此串中额外的 t 个 0 只是领着我们沿着那个循环多走一次，并将 s_i 再带回它自己（如

图 11-19所示)。然后，此串的剩余部分所导致的状态与以前完全相同。这个矛盾证明了$\{0^n1^n \mid n=1, 2, \cdots\}$不是正则的。■

11.4.6 一些更强大的机器

很多计算都不能用有限状态自动机来完成，这类机器的局限性是它们只有有限的存储，因而限制了它们识别那些不是正则的语言，如$\{0^n1^n \mid n=0, 1, 2, \cdots\}$。因为一个集合是正则的当且仅当它是一个正则文法生成的语言，例 5 证明了没有正则文法能够生成集合$\{0^n1^n \mid n=0, 1, 2, \cdots\}$。但是，有一个上下文无关文法能够识别这个集合，此文法就是 11.1 节例 5 所给的文法。

由于有限状态自动机的局限性，有必要使用其他更加强大的计算模型。**下推自动机**就是这样一个模型。下推自动机除了包括有限状态自动机的所有部件外，还有一个栈，此栈能提供无限的存储。可以将符号放在栈顶上，也可从栈顶提走符号。下推自动机以两种方式识别集合。其一是，如果一个集合是所有这样的串构成的，当它们作为输入时产生空栈，则此集合能被识别。其二是，如果一个集合是所有这样的串构成的，当它们作为输入时导致终结状态，则此集合能被识别。可以证明，一个集合能被下推自动机识别，当且仅当它是一个上下文无关文法生成的语言。

但是，还有一些集合不能表示成上下文无关文法生成的语言，如集合$\{0^n1^n2^n \mid n=0, 1, 2, \cdots\}$。我们将指出为什么这个集合不能被下推自动机识别，但不给出证明，因为还没有介绍所需的方法。(但本章末尾的补充练习 28 给出了一个证明方法。)可以使用栈来查看一个串是否以一列 0 开始，后再跟相同个数的 1，做法是对每个 0(只要仅读到多个 0 时)在栈上放一个符号，对每个 1(只要仅读到 0 后面的多个 1 时)从栈中去掉一个这样符号。但这个过程一旦完成，栈就空了，也就没法判断此串中是否还有与 0 个数相同的一列 2。

还有一种比下推自动机更强大的机器，叫**线性有界自动机**，它能识别诸如$\{0^n1^n2^n \mid n=0, 1, 2, \cdots\}$的集合。特别地，线性有界自动机能够识别上下文相关文法。但是，这些机器不能识别短语结构文法生成的所有语言。为避免特殊类型机器的局限性，我们使用一种称为**图灵机**的模型，这种机器是以英国数学家图灵[⊖]命名的。图灵机除了包括有限状态自动机的所有部件外，还有一个带，其两端都是无限的。图灵机具有在带上读和写的能力，还能沿着带子左右移动。图灵机能够识别短语结构文法生成的所有语言。另外，它还能作为在计算机器上执行的所有计算



⊖ **图灵(Alan Mathison Turing，1912—1954)** 图灵是父亲在印度民政部供职时孕于母腹，但在伦敦出生。他在孩提时代就对化学和机械着迷，做过大量化学实验。图灵曾就读于英国的寄宿学校 Sherborne。1931 年，他获得了剑桥大学皇家学院的奖学金。在完成毕业论文后，他被选为该学院的成员。在毕业论文中，他重新发现了统计学中的一个著名定理——中心极限定理。1935 年，他对判定问题着了迷，这是伟大的德国数学家希尔伯特提出的一个问题。问题是：是否有一个能用于判断任何命题是否为真的一般方法。图灵喜欢跑步(后来业余跑步成为他生活中酷爱的一项活动)。一天，在跑步之后的休息中，他发现了解决判定问题的关键思想。在他的解决方案中，他发明了现今称为图灵机的东西，并用它作为计算机器的最一般模型。利用这个机器，他发现了一个不能用一般方法判定的问题，这是一个关于他称为可计算数的问题。

从 1936 到 1938 年，图灵在普林斯顿大学访问，与丘奇(Alonzo Church)一起工作。丘奇也解决了希尔伯特的判定问题。在 1939 年，图灵回到了皇家学院。但在第二次世界大战爆发期间，他进入了英国外交部，从事对德国密码的分析工作。他对破解机械的德国密码机 Enigma 的密码做出了重要贡献，在赢得这次战争中起到了重要作用。

战后，图灵从事早期计算机的开发。他对机器的思考能力产生了兴趣。他认为如果一台计算机在对问题的书面答复中与人没有区别，则应该认为它在"思考"。他还对生物学感兴趣，曾经写过关于有机体形式的形成和发展的书。1954 年，图灵服氰化物自杀，没有留下遗言作明确解释。也许，由于涉及同性恋关系的法律困扰和法院强迫他进行荷尔蒙治疗以减少性冲动，促使他决定结束自己的生命。

的模型。由于这个能力，图灵机在理论计算机科学中得到了广泛得研究。下一节将作简要介绍。

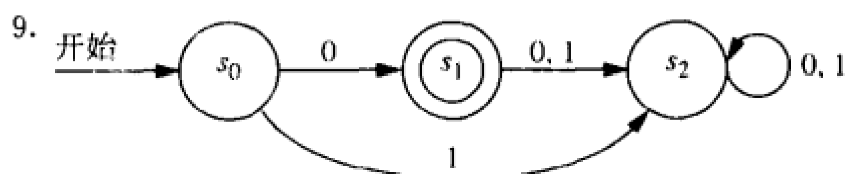
练习

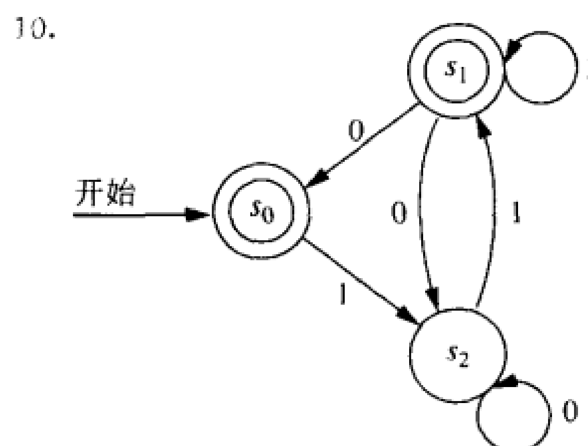
1. 用文字描述下列每个正则集合中的串。
   a) $\mathbf{1^*0}$　b) $\mathbf{1^*00^*}$　c) $\mathbf{111 \cup 001}$
   d) $\mathbf{(1 \cup 00)^*}$　e) $\mathbf{(00^*1)^*}$　f) $\mathbf{(0 \cup 1)(0 \cup 1)^*00}$
2. 判断 1011 是否属于下列正则集合。
   a) $\mathbf{10^*1^*}$　b) $\mathbf{0^*(10 \cup 11)^*}$　c) $\mathbf{1(01)^*1^*}$
   d) $\mathbf{1^*01(0 \cup 1)}$　e) $\mathbf{(10)^*(11)^*}$　f) $\mathbf{1(00)^*(11)^*}$
   g) $\mathbf{(10)^*1011}$　h) $\mathbf{(1 \cup 00)(01 \cup 0)1^*}$
3. 用正则表达式表达下列每个集合。
   a)一个或更多的 0 后面跟一个 1 形成的串的集合。
   b)两个或两个以上符号后面跟 3 个或 3 个以上 0 形成的串的集合。
   c)一个 0 前没有 1 或一个 1 前没有 0 的串的集合。
   d)集合包含这样的串：先是个数为 2(mod 3)的一串 1，后面是偶数个 0。
4. 构造确定型有限状态自动机来识别下列包含在 I^* 中的集合(其中 I 是一个字母表)。
   a) $\emptyset$　b) $\{\lambda\}$　c) $\{a\}$，其中 $a \in I$

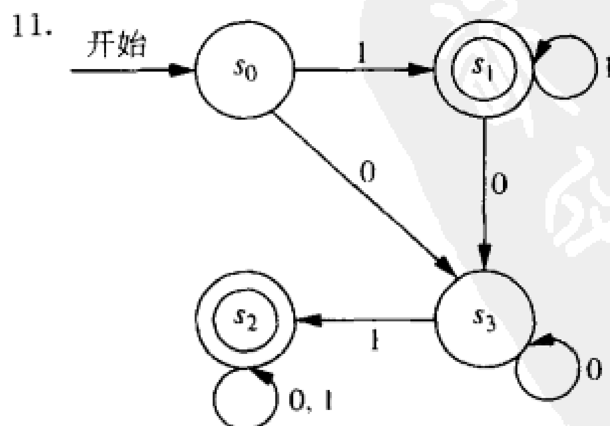
*5. 若 A 是一个正则集合。证明：A 中串的反串构成的集合 A^R 也是正则的。

6. 求识别下列集合的有限状态自动机。
   a) $\{\lambda, 0\}$　b) $\{0, 11\}$　c) $\{0, 11, 000\}$
7. 用克莱因定理的证明中描述的构造方法，求识别下列集合的非确定型有限状态自动机。
   a) $\mathbf{0^*1^*}$　b) $\mathbf{(0 \cup 11)^*}$　c) $\mathbf{01^* \cup 00^*1}$
8. 构造非确定型有限状态自动机，用它识别正则文法 $G=(V, T, S, P)$生成的语言，其中：$V=\{0, 1, S, A, B\}$，$T=\{0, 1\}$，S 是初始符号，产生式集合为
   a) $S \to 0A$，$S \to 1B$，$A \to 0$，$B \to 0$
   b) $S \to 1A$，$S \to 0$，$S \to \lambda$，$A \to 0B$，$B \to 1B$，$B \to 1$
   c) $S \to 1B$，$S \to 0$，$A \to 1A$，$A \to 0B$，$A \to 1$，$A \to 0$，$B \to 1$

在练习 9～11 中，构造正则文法 $G=(V, T, S, P)$，使之生成的语言是所给的有限状态机识别的语言。

9.


10.


11.


12. 在定理2的证明中，从正则文法构造了一个有限状态自动机。证明：此自动机识别这个文法生成的集合。
13. 在定理2的证明中，从正则文法构造了一个有限状态自动机。证明：此文法生成这个自动机识别的集合。
14. 证明每个非确定型有限状态自动机等价于另一个这样的非确定型有限状态自动机，它的初始状态永不会被再次访问。
*15. 设 $M=(S, I, f, s_0, F)$ 是一个确定型有限状态自动机。证明：M 识别的语言 $L(M)$ 是无限的，当且仅当存在一个能被 M 识别的词 x 满足 $l(x)\geqslant|S|$。
*16. 用来证明某个集合不是正则的一个重要技术是泵引理。泵引理表述为：如果 $M=(S, I, f, s_0, F)$ 是一个确定型有限状态自动机，x 是 M 识别的语言 $L(M)$ 中的一个串，$l(x)\geqslant|S|$。则存在 I^* 中的串 u，v 和 w，使得 $x=uvw$，$l(uv)\leqslant|S|$，$l(v)\geqslant 1$，且 $uv^iw\in L(M)(i=0, 1, 2, \cdots)$。证明泵引理。[提示：使用例5中的思想。]
*17. 证明集合 $\{0^{2n}1^n \mid n=0, 1, 2, \cdots\}$ 不是正则的。可以使用练习16中的泵引理。
*18. 证明集合 $\{1^{n^2} \mid n=0, 1, 2, \cdots\}$ 不是正则的。可以使用练习16中的泵引理。
*19. 证明 $\{0, 1\}$ 上所有回文构成的集合不是正则的。可以使用练习16中的泵引理。[提示：考察形如 0^N10^N 的串。]
**20. 证明被有限状态自动机识别的集合是正则的。(这是克莱因定理的充分性部分。)

11.5 图灵机

11.5.1 引言

本章前面部分研究的有限状态自动机不能作为计算的一般模型，因为它们有其自身的局限性。例如，有限状态自动机虽然能识别正则集合，却不能识别许多很容易描述的集合，如 $\{0^n1^n \mid n\geqslant 0\}$，计算机使用存储才能识别这些集合。可以用有限状态自动机来计算一些相对简单的函数(如两个数的和)，但不能用它们来计算计算机所计算的函数(如两个数的积)。为克服这些不足，我们使用一种更强大的机器，称为图灵机，它是以著名数学家和计算机科学家图灵(Alan Turing)的名字命名的，他在20世纪30年代发明了这种机器。

图灵机主要由一个控制头和一个带组成，控制头在任何时候都处于有限多个不同状态中的某个状态，带子被分成许多方格，且两端都是无限的。当图灵机的控制头沿着带子来回移动时，他能够在带上读和写，并根据所读的带符号改变状态。图灵机比有限状态自动机更强大，因为他有存储能力，而有限状态自动机却没有。我们将说明怎么用图灵机来识别集合，包括识别有限状态自动机不能识别的集合。还将说明怎么用图灵机来计算函数。图灵机是计算的最一般模型，本质上，它能做计算机能做的任何事情。

11.5.2 图灵机的定义

下面给出图灵机的形式定义。之后将根据它的控制头的动作来解释这个形式定义，控制头的动作包括读或写带上的符号及沿着带子左右移动。

定义1 图灵机 $T=(S, I, f, s_0)$ 由下列各部分组成：有限状态集 S，包含空白符 B 的字母表 I，从 $S\times I$ 到 $S\times I\times\{R, L\}$ 的部分函数 f，及初始状态 s_0。

回忆1.8节练习69的引言，部分函数只对定义域中的某些元素有定义。这意味着上述部分函数 f 对于某些(状态，符号)对没有定义。但对于有定义的对，只有唯一一个三元组(状态，符

号，方向)与之对应。

为用机器的观点来解释这个定义，考察控制头和带。如图 11-20 所示，带被分成小方格，且两端都是无限的，在任何时刻其上都只有有限多个非空白符。图灵机运行的每一步动作依赖于部分函数对当前状态和带符号的值。

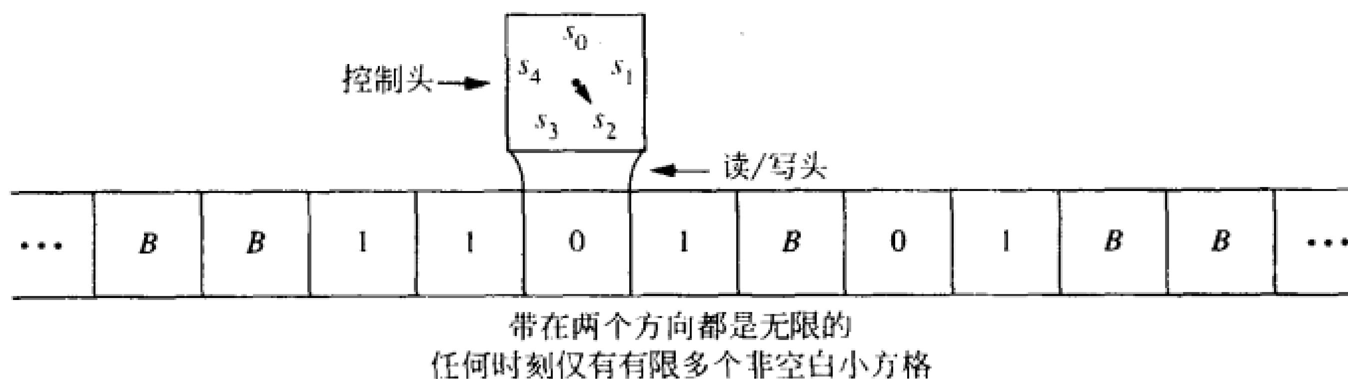


图 11-20 图灵机的表示

在每一步，控制头读的是当前带符号 x。如果控制头处于状态 s，且部分函数 f 在(s, x)处由 $f(s, x)=(s', x', d)$定义，则控制头

1)进入状态 s'。

2)在当前方格中擦掉 x，并写上符号 x'。

3)如果 $d=R$，向右移动一个方格；如果 $d=L$，向左移动一个方格。

我们将这一步写成五元组(s, x, s', x', d)。如果部分函数 f 在(s, x)处没有定义，则图灵机 T 就停机。

定义一个图灵机的常用方法是指明形如(s, x, s', x', d)的五元组构成的一个集合。当使用这个定义时，就隐式定义了状态集和输入字母表。

在运行开始的时候，总假设图灵机处于初始状态 s_0，且处于带中最左边的非空白符上。如果带中都是空白符，则控制头可以处于任何方格上。控制头所在的最左边非空白符位置称为该机器的**初始位置**。

下面的例子说明图灵机是怎么运行的。

例 1 下列 7 个五元组定义一个图灵机 T：$(s_0, 0, s_0, 0, R)$，$(s_0, 1, s_1, 1, R)$，(s_0, B, s_3, B, R)，$(s_1, 0, s_0, 0, R)$，$(s_1, 1, s_2, 0, L)$，(s_1, B, s_3, B, R)，$(s_2, 1, s_3, 0, R)$。当 T 在图 11-21a 所示的带上运行时，最后的带子是什么？

解 在开始运行时，T 处于状态 s_0，且在带中最左边的非空白符上。第一步，根据五元组$(s_0, 0, s_0, 0, R)$，读最左边的非空白方格中的 0，保持状态 s_0，在此方格中写下 0，向右移动一个方格。第二步，根据五元组$(s_0, 1, s_1, 1, R)$，读当前方格中的 1，进入状态 s_1，在这方格中写下 1，向右移动一个方格。第三步，根据五元组$(s_1, 0, s_0, 0, R)$，读当前方格中的 0，进入状态 s_0，在这方格中写下 0，向右移动一个方格。第四步，根据五元组$(s_0, 1, s_1, 1, R)$，读当前方格中的 1，进入状态 s_1，在这方格中写下 1，向右移动一个方格。第五步，根据五元组$(s_1, 1, s_2, 0, L)$，读当前方格中的 1，进入状态 s_2，在这方格中写下 0，向左移动一个方格。第六步，根据五元组$(s_2, 1, s_3, 0, R)$，读当前方格中的 1，进入状态 s_3，在这方格中写下 0，向右移动一个方格。最后，机器在第七步停机，因为在这个机器的描述中，没有五元组是以$(s_3, 0)$开头的。所有这些步骤如图 11-21 所示。

注意，T 将带上第一对连续的 1 变为 0 后停机。 ■

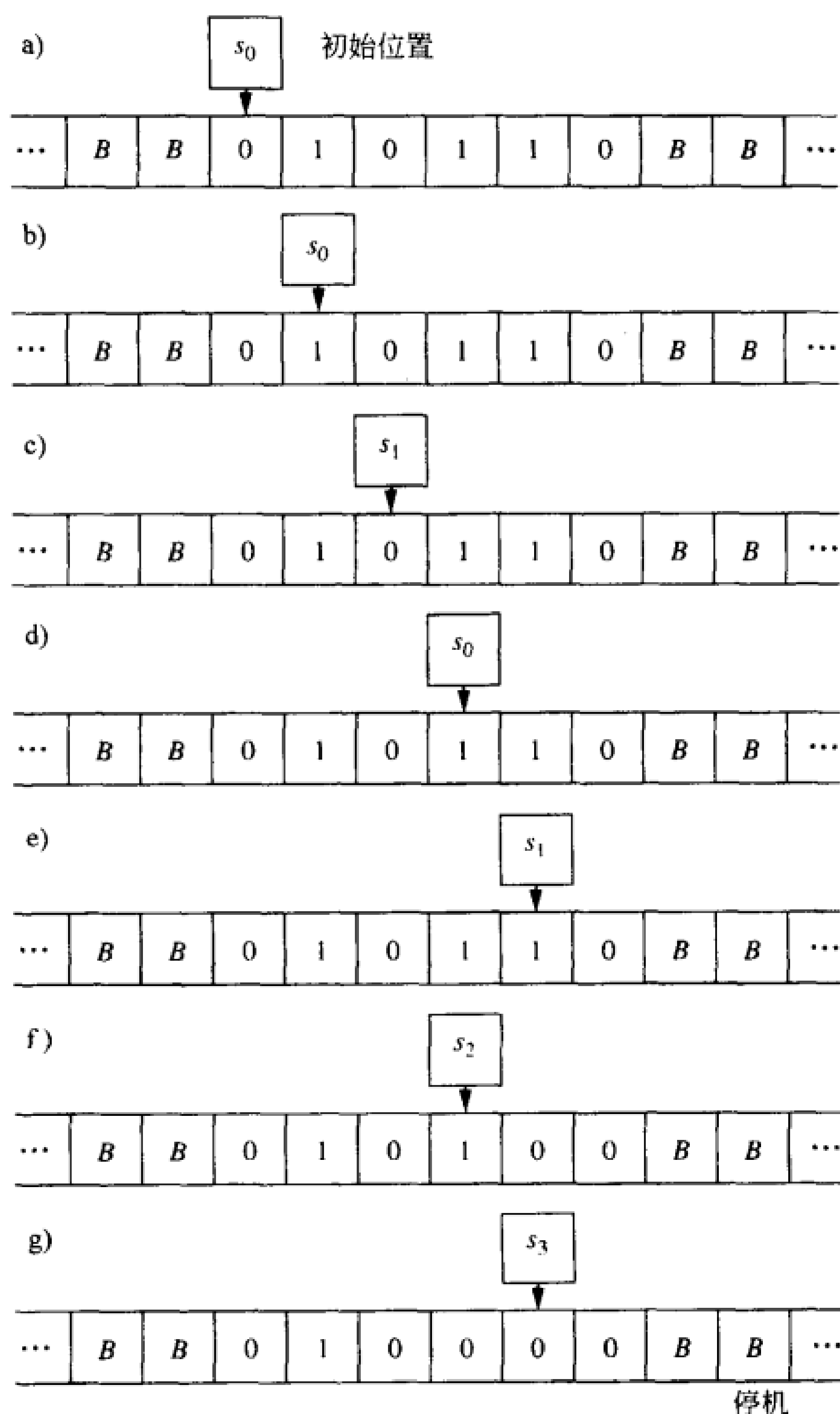


图 11-21 图灵机 T 在图 11-20 所示带上的运行步骤

11.5.3 用图灵机识别集合

可以用图灵机来识别集合。为此，如下定义终结状态的概念。图灵机 T 的终结状态是这样的状态：在描述 T 的五元组中，此状态不是任何五元组的第一个状态(例如，例 1 中的状态 s_3)。

现在定义图灵机识别一个串的含义是什么。给定一个串，我们在连续的方格中写下此串中的连续的符号。

定义 2 设 V 是字母表 I 的一个子集。图灵机 $T=(S, I, f, s_0)$ 识别 V^* 中串 x 当且仅当：若将 x 写在带上，T 从初始位置开始运行，则 T 能在一个终结状态停机。称 T 能识别 V^* 的子集 A，如果 x 能被 T 识别当且仅当 x 属于 A。

注意，为了识别 V^* 的子集 A，我们可以使用不在 V 中的符号。也就是说，输入字母表 I 也许包含不是 V 中的符号。这些额外的符号常用来当作标记(见例 3)。

什么情况下图灵机 T 不识别 V^* 中串 x 呢？答案是：设 x 的符号被放在 T 的带上的连续方格中，T 从初始位置开始运行，若 T 不停机，或者虽然停机，但不在终结状态停机，则 T 不识别 x。（读者应该明白，这是定义图灵机如何识别集合的许多方法中的一种。）

下面的例子说明了这个概念。

例 2 求一个图灵机，使之能识别第二位是 1 的二进制串构成的集合（即正则集合 $(\mathbf{0}\cup\mathbf{1})\mathbf{1}(\mathbf{0}\cup\mathbf{1})^*$）。

解 我们想要如下的图灵机：它从最左边的非空白带方格开始运行，然后向右移动，并且判断第二个符号是否为 1。若第二个符号是 1，则机器应该进入终结状态；如果第二个不是 1，则机器不能停机，或者在一个非终结状态下停机。

为构造这样的图灵机，应该包括五元组 $(s_0, 0, s_1, 0, R)$ 和 $(s_0, 1, s_1, 1, R)$ 来读第一个符号，并进入状态 s_1。下一步，添加五元组 $(s_1, 0, s_2, 0, R)$ 和 $(s_1, 1, s_3, 1, R)$ 来读第二个符号，且当这个符号是 0 时，进入状态 s_2。当这个符号是 1 时，进入状态 s_3。我们不希望使第二位是 0 的串也被识别，所以 s_2 不能是终结状态。我们希望 s_3 是终结状态。所以我们要加入五元组 $(s_2, 0, s_2, 0, R)$。因为我们不想识别空串和只有一位的串，所以还加入五元组 $(s_0, B, s_2, 0, R)$ 和 $(s_1, B, s_2, 0, R)$。

由上述 7 个五元组组成的图灵机 T 在终结状态 s_3 终止当且仅当：此二进制输入串至少有 2 位，并且第二位是 1。如果此二进制串少于两位，或着其第二位不是 1，则机器将在非终结状态 s_2 终止。 ■

对于一个正则集合，可以构造一个总是向右移动的图灵机来识别它（如例 2）。为构造这样的图灵机，先构造一个识别此集合的有限状态自动机，然后再用此有限状态自动机的转移函数构造一个图灵机，使之总向右移动。

下面说明怎么构造图灵机来识别非正则集合。

例 3 求识别集合 $\{0^n1^n \mid n\geq 1\}$ 的图灵机。

解 为构造这样图灵机，我们使用辅助带符号 M 作为标记。令 $V=\{0, 1\}$，$I=\{0, 1, M, B\}$。我们希望只识别 V^* 中的串。我们还有一个终结状态 s_6。图灵机依次用 M 替换串中最左边的 0，和用 M 替换串中最右边的 1，这样在带上左右移动。它能在一个终结状态终止当且仅当：这个串的构成是一列 0 后跟一列相同个数的 1。

虽然这很容易描述，图灵机也很容易执行，但我们想要使用的图灵机本身却有点复杂。标记 M 是用来跟踪已经检查过的最左边和最右边的符号。所用的五元组是：$(s_0, 0, s_1, M, R)$，$(s_1, 0, s_1, 0, R)$，$(s_1, 1, s_1, 1, R)$，(s_1, M, s_2, M, L)，(s_1, B, s_2, B, L)，$(s_2, 1, s_3, M, L)$，$(s_3, 1, s_3, 1, L)$，$(s_3, 0, s_4, 0, L)$，(s_3, M, s_5, M, R)，$(s_4, 0, s_4, 0, L)$，(s_4, M, s_0, M, R)，(s_5, M, s_6, M, R)。例如，当机器从开始一直运行到停机时，串 000111 将依次变成 $M00111$，$M0011M$，$MM011M$，$MM01MM$，$MMM1MM$，$MMMMMM$。（注意，此串不是根据图灵机的操作步骤变化的。）

解释这个图灵机的动作和它为什么能识别集合 $\{0^n1^n \mid n\geq 1\}$ 将留给读者作为练习（本节末的练习 13）。 ■

可以证明：一个集合能被图灵机识别当且仅当它是 0 型文法生成的集合，即短语结构文法生成的集合。这里略去它的证明。

11.5.4 用图灵机计算函数

图灵机可以看作是能求部分函数的值的计算机。为理解这一点，假设给定输入串 x 时图灵机 T 能够停机，且停机时，串 y 在它的带上。因此可以定义 $T(x)=y$。T 的定义域是使 T 能停

机的串构成的集合。对于输入 x，若 T 不停机，则 $T(x)$没有定义。将图灵机看成计算串的函数值的机器是有用的，但怎么用图灵机来计算定义在整数、整数对或整数三元组等上的函数呢？

为将图灵机看作是计算 k 元非负整数集合到非负整数集合的函数(这样的函数称为*数论函数*)的计算机，需要找到在带上表示 k 元整数的方法。为此，我们使用整数的*一元表示*，即将非负整数 n 表示成有 $n+1$ 个 1 的串。例如，0 被表示成串 1，5 被表示成串 111111。为表示 k 元组(n_1，n_2，…，n_k)，我们先写 n_1+1 个 1，后面跟一个星号，再跟 n_2+1 个 1，再跟一个星号，等等，以 n_k+1 个 1 结尾。例如，四元组(2，0，1，3)可以表示成串 111 * 1 * 11 * 1111。

现在能将图灵机看成计算一系列数论函数 T，T^2，…，T^k，…。函数 T^k 是根据 T 在 k 元整数组上的动作定义的，当然，k 元整数组被表示成用星号隔开的一些一元表示。

例 4 构造一个图灵机将两个非负整数相加。

解 需要构造图灵机来计算函数 $f(n_1, n_2)=n_1+n_2$。元素对(n_1，n_2)被表示成这样的串：先是 n_1+1 个 1，后面跟一个星号，再跟 n_2+1 个 1。机器 T 应以这个串作为输入，并在带上产生 n_1+n_2+1 个 1 作为输出。实现这个任务的一个方法如下：机器从输入串最左边的 1 开始运行，执行去掉这个 1 的步骤。若 $n_1=0$，则停机，此时，星号之前已没有 1 了。在剩下的 1 中，以最左边的 1 替换星号，然后停机。下列五元组能做到这一点：(s_0，1，s_1，B，R)，(s_1，*，s_3，B，R)，(s_1，1，s_2，B，R)，(s_2，1，s_2，1，R)，(s_2，*，s_3，1，R)。 ■

不幸的是，即使是相对简单的函数，要构造图灵机来计算它也是极为费力的。例如，在许多书中都有计算两个非负整数乘积的图灵机，此图灵机有 31 个五元组和 11 个状态。如果构造计算相对简单的函数的图灵机都是挑战性的，那么我们对构造更加复杂函数的图灵机还有什么指望呢？简化这个问题的一个方法是使用多带图灵机(它同时使用不止一个带子)，并给出构造复合函数的多带图灵机的方法。可以证明：对任何多带图灵机，存在一个单带图灵机，使得他们能做完全相同的事情。


可由图灵机计算的函数称为*可计算的*。可以很直接地证明：存在不是可计算的数论函数，但要实际给出这样一个函数却不容易。在本节末尾的练习 23 中，其导语定义了一个*忙碌海狸函数*，这是一个不可计算函数。证明忙碌海狸函数是不可计算的一种方法是证明它比任何可计算函数都增长得快(见练习 24)。

11.5.5 不同类型的图灵机

图灵机的定义有许多变种。可用很多方法来扩展图灵机的能力。例如，可以允许图灵机在一步中左称、右移或根本不动；允许图灵机操作多个带子，n 个带的图灵机可以用$(2+3n)$元组来描述；允许带是二维的，即在每一步可以上下左右移动，而不像在一维带上那样只向左或右移动；还可以允许有多个带头，它们能同时读不同的方格。还有，可以允许图灵机是非确定的，即允许(状态，带符号)对作为头两个元素出现在图灵机的多个五元组中。也可以用多种方法来削减图灵机的能力。例如，可以限制带只在一个方向是无限的；可以限制带字母表只有两个符号。图灵机的所有这些变种都已被详细地研究。

重要的是：不管使用哪个变种图灵机，或使用变种图灵机的哪个组合，都决不会增加或减少机器的能力。这些变种的任何一个能做的事，本节定义的图灵机都能做到，反之亦然。这些变种之所以还有用，是因为：有些时候，在做某些特殊任务时，使用它们比只使用定义 1 定义的图灵机容易得多。它们永远不会扩展机器的能力。

11.5.6 丘奇-图灵论题


图灵机还是相对简单的。它只能有有限多个状态，每一次它们只能在一维带上读或写一个符号。但结果表明，图灵机是极其强大的，我们已经看到，可以构造图灵机来

执行数的加法和乘法。对于能够用一个算法计算的特殊函数，虽然很难实际构造图灵机来计算它们，但这样的图灵机总是能够找到的。这也正是图灵发现这种机器的目的。

*丘奇-图灵论题*是说：对于任何可用有效算法来解的问题，都存在解该问题的图灵机。可以用大量的证据来说明丘奇-图灵论题，但它还是被称为是论题，而不是定理，这是因为：有效算法说明的可解性概念是非形式的且不严格的，相反，图灵机定义的可解性概念是形式的且是严格的。当然，对于任何问题，只要它能够用带有用某个语言写成的程序的计算机来解，即使使用了无限多的存储，都应该被认为是有效可解的⊖。

人们发明了许多形式理论来刻画有效可计算性概念，其中有图灵的理论、丘奇⊖的λ演算以及克莱因和波斯特(Post)提出的理论。这些理论表面上看起来十分不同，但令人惊奇的是，它们都是等价的，因为可以证明它们定义了完全相同的函数类。由此证据可以看出，图灵的思想虽然是在现代计算机的发明之前形成的，但确实描述了计算机最根本的能力。

练习

1. 设 T 是下列五元组定义的图灵机：$(s_0, 0, s_1, 1, R)$，$(s_0, 1, s_1, 0, R)$，$(s_0, B, s_1, 0, R)$，$(s_1, 0, s_2, 1, L)$，$(s_1, 1, s_1, 0, R)$，$(s_1, B, s_2, 0, L)$。对于下列初始带，判断 T 停机时的最终带。假设 T 从初始位置开始执行。

a)

…	B	B	0	0	1	1	B	B	…

b)

…	B	B	1	0	1	B	B	B	…

c)

…	B	B	1	1	B	0	1	B	…

d)

…	B	B	B	B	B	B	B	B	…

2. 设 T 是下列五元组定义的图灵机：$(s_0, 0, s_1, 0, R)$，$(s_0, 1, s_1, 0, L)$，$(s_0, B, s_1, 1, R)$，$(s_1, 0, s_2, 1, R)$，$(s_1, 1, s_1, 1, R)$，$(s_1, B, s_2, 0, R)$，$(s_2, B, s_3, 0, R)$。对于下列初始带，判断 T 停机时的最终带。假设 T 从初始位置开始执行。

a)

…	B	B	0	1	0	1	B	B	…

b)

…	B	B	1	1	1	B	B	B	…

c)

…	B	B	0	0	B	0	0	B	…

d)

…	B	B	B	B	B	B	B	B	…

3. 对于由五元组$(s_0, 0, s_0, 0, R)$，$(s_0, 1, s_1, 0, R)$，(s_0, B, s_2, B, R)，$(s_1, 0, s_1, 0, R)$，$(s_1, 1, s_0, 1, R)$和(s_1, B, s_2, B, R)描述的图灵机，给定一个二进位串输入，它能做什么？
4. 对于由五元组$(s_0, 0, s_1, B, R)$，$(s_0, 1, s_1, 1, R)$，$(s_1, 0, s_1, 0, R)$，$(s_1, 1, s_2, 1, R)$，$(s_2, 0, s_1, 0, R)$，$(s_2, 1, s_3, 0, L)$，$(s_3, 0, s_4, 0, R)$和$(s_3, 1, s_4, 0, R)$描述的图灵机，给定一个二进位串输入，它能做什么？
5. 构造一个带符号为 0、1 和 B 的图灵机，它将带上第一个 0 替换为 1，而其余符号保持不变。

⊖ "有效可解"在有些参考书中又称为"能行可解"。——译者注

⊖ **丘奇(Alonzo Church, 1903—1995)** 丘奇出生于华盛顿特区，曾在哥廷根跟随希尔伯特学习，后来转到阿姆斯特丹。从 1927 年到 1967 年，他执教于普林斯顿大学，1967 年调到加州大学洛杉矶分校(UCLA)。丘奇是符号逻辑学会(Association of Symbolic Logic)的创始人之一。他对可计算性理论作出了实质性的贡献，其中包括对判定问题的解、λ演算的发明，以及对现今称为丘奇-图灵论题的陈述。克莱因和图灵都是丘奇的学生。他在 90 岁生日后还在发表文章。

6. 构造一个带符号为 0、1 和 B 的图灵机，对于给定的二进制输入串，它将带上所有 0 替换为 1，而所有的 1 保持不变。
7. 构造一个带符号为 0、1 和 B 的图灵机，对于给定的二进制输入串，它将带上最左边的 1 以外的所有 1 替换为 0，而其余符号保持不变。
8. 构造一个带符号为 0、1 和 B 的图灵机，对于给定的二进制输入串，它将带上首先出现的两个连续的 1 替换为 0，而其余符号保持不变。
9. 构造一个图灵机，它识别的集合是所有以 0 结尾的二进制串组成的集合。
10. 构造一个图灵机，它识别的集合是所有至少包含两个 1 的二进制串组成的集合。
11. 构造一个图灵机，它识别的集合是所有包含偶数个 1 的二进制串组成的集合。
12. 对于例 3 中的图灵机，若从下列每个串开始运行，写出其每一步的带内容。
    a)0011　　b)00011　　c)101100　　d)000111
13. 例 3 中的图灵机识别一个串当且仅当此串具有形式 0^n1^n（其中 n 是一个正整数），试说明原因。

*14. 构造识别集合 $\{0^{2n}1^n \mid n\geqslant 0\}$ 的图灵机。

*15. 构造识别集合 $\{0^n1^n2^n \mid n\geqslant 0\}$ 的图灵机。

16. 构造一个图灵机计算函数 $f(n)=n+2$，其中 n 是非负整数。
17. 构造一个图灵机计算下列函数：当 $n\geqslant 3$ 时，$f(n)=n-3$，当 $n=0$，1，2 时，$f(n)=0$，其中 n 是非负整数。
18. 构造一个图灵机计算函数 $f(n)=n \bmod 3$。
19. 构造一个图灵机计算下列函数：当 $n\geqslant 5$ 时，$f(n)=3$；当 $n=0$，1，2，3 或 4 时，$f(n)=0$。
20. 构造一个图灵机计算下列函数：对于所有非负整数 n_1 和 n_2，$f(n_1, n_2)=n_2+2$。

*21. 构造一个图灵机计算下列函数：对于所有非负整数 n_1 和 n_2，$f(n_1, n_2)=\min\{n_1, n_2\}$

22. 构造一个图灵机计算下列函数：对于所有非负整数 n_1 和 n_2，$f(n_1, n_2)=n_1+n_2+1$。

设 $B(n)$ 是具有 n 个状态且字母表为 $\{1, B\}$ 的图灵机从空白带开始运行后在带上所能打印的 1 的最大个数。根据给定的值 n 确定 $B(n)$ 这个问题称为*忙碌海狸问题*(busy beaver problem)，该问题由拉多(Tibor Rado)于 1962 年首先研究。现在已经知道：$B(2)=4$，$B(3)=6$，$B(4)=13$。但当 $n\geqslant 5$ 时，$B(n)$ 等于什么还不知道。

*23. 通过寻找下面的图灵机来证明 $B(2)$ 至少是 4：该图灵机有两个状态，字母表是 $\{1, B\}$，且在停机时，带上有 4 个连续的 1。

**24. 证明函数 $B(n)$ 不能用任何图灵机来计算。[*提示*：假设有一个图灵机计算二进制表示的 $B(n)$，构造一个图灵机 T，从空带开始，写下 n 的二进制表示，计算 $B(n)$ 并表示成二进制数，然后将 $B(n)$ 从二进制表示转换为一元表示。再证明：当 n 充分大时，T 的状态个数可以小于 $B(n)$，导致矛盾。]

关键术语和结果

术语

字母表(或词汇表)：用来构造串的元素组成的集合。

语言：字母表上所有串构成的集合的一个子集。

短语结构文法(V, T, S, P)：语言的一种描述，包括字母表 V、终结符集 T、初始符号 S 和产生式集 P。

产生式 $w\rightarrow w_1$：只要语言的某个串中出现了 w，就可将此串中的 w 替换为 w_1。

$w_1\Rightarrow w_2$(*由 w_1 直接派生 w_2*)：w_2 是从 w_1 按如下方式得到的：用产生式将 w_1 中的某个串替换为另一个串。

$w_1 \stackrel{*}{\Rightarrow} w_2$（由 w_1 直接派生 w_2）：w_2 是从 w_1 按如下方式得到的：用一系列产生式将某些串替换为另一些串。

0 型文法：任意短语结构文法。

1 型文法：是一种短语结构文法，但其产生式都具有形式 $w_1 \to w_2$，其中 $l(w_1) \leqslant l(w_2)$ 或 $w_2 = \lambda$。

2 型(或上下文无关)文法：是一种短语结构文法，但其产生式都具有形式 $A \to w_1$，其中 A 是一个非终结符。

3 型(或正则)文法：是一种短语结构文法，但其产生式的形式是 $A \to aB$，$A \to a$ 或 $S \to \lambda$，其中 A，B 是非终结符，S 是初始符，a 是一个终结符。

派生(或语法分析)树：一个带根的有序树，其根表示 2 型文法的初始符，内部结点表示非终结符，叶表示终结符，结点的儿子是产生式右边的符号，按从左到右顺序排列，同时，父亲表示的符号都在左边。

巴克斯-诺尔范式：上下文无关文法的一种描述，在这种描述中，左边非终结符相同的所有产生式被合并成一个式子，式子的右边是这些产生式不同的右边，并用竖线符将其分开，用尖括号将非终结符括起来，符号→被换成∷=。

有限状态机器(S，I，O，f，g，s_0)(或米利机)：一个六元组，包括状态集 S，输入字母表 I，输出字母表 O，转移函数 f(对每个状态与输入对，指派下一个状态)，输出函数 g(对每个状态与输入对，指派一个输出)和一个初始符 s_0。

AB(A 和 B 的连接)：由 A 中串和 B 中串连接而成的串构成的集合。

A^*(A 的克莱因闭包)：由 A 中任意多个串连接而成的串构成的集合。

确定型有限状态自动机(S，I，f，s_0，F)：一个五元组，包括状态集 S，输入字母表 I，转移函数 f(对每个状态与输入对，指派下一个状态)，初始符 s_0 和终结状态集 F。

非确定型有限状态自动机(S，I，f，s_0，F)：一个五元组，包括状态集 S，输入字母表 I，转移函数 f(对每个状态与输入对，指派下一个可能状态的集合)，初始符 s_0 和终结状态集 F。

自动机识别的语言：将自动机从初始状态带到终结状态的输入串构成的集合。

正则表达式：如下递归定义的表达式：$\varnothing$，λ 和每个输入符号 x 都是正则表达式；当 A 和 B 是正则表达式时，(AB)，(AB)和 A^* 都是正则表达式。

正则集合：正则表达式定义的集合。

图灵机 $T=(S, I, f, s_0)$：由下列各部分组成的四元组：有限状态集 S，包含空白符 B 的字母表 I，从 $S \times I$ 到 $S \times I \times \{R, L\}$ 的一个部分函数，初始状态 s_0。

结果

对每个非确定型有限状态自动机，存在一个确定型有限状态自动机，它们识别相同的集合。

克莱因定理：一个集合是正则的，当且仅当它可由一个有限状态自动机来识别。

一个集合是正则的当且仅当它可由一个正则文法生成。

复习题

1. a)定义短语结构文法。

   b)“一个串可以由短语结构文法从串 w 派生出来”的含义是什么？

2. a)什么是短语结构文法生成的语言？

   b)设短语结构文法 G 如下：词汇表为{S，0，1}，终结符集为 T={0，1}，初始符号为 S，产生式为 $S \to$

000S 和 $S\to 1$。G 生成的语言是什么？

c)给出生成集合$\{01^n \mid n=0, 1, 2, \cdots\}$的短语结构文法。

3. a)定义一个 1 型文法。
   b)给出一个是文法却不是 1 型文法的例子。
   c)定义 2 型文法。
   d)给出一个是 1 型文法却不是 2 型文法的例子。
   e)定义一个 3 型文法(或上下文无关文法)。
   f)给出一个是 2 型文法却不是 3 型文法的例子。
4. a)定义一个正则文法。
   b)定义一个正则语言。
   c)给出一个是 3 型文法却不是正则文法的例子
   d)证明集合$\{0^m 1^n \mid m、n=0, 1, 2, \cdots\}$是一个正则语言。
5. a)什么是巴克斯-诺尔范式？
   b)选择英语的一个子集，给出其巴克斯-诺尔范式。
6. a)什么是一个有限状态机？
   b)怎么用有限状态机建立下列自动售货机的模型：它只接受 15 分硬币，在放入 75 分之后，它发售一筒软饮料。
7. a)什么是一个串集合的克莱因闭包？
   b)求集合$\{11, 0\}$的克莱因闭包。
8. a)定义一个有限状态自动机。
   b)“一个串由一个有限状态自动机识别”的含义是什么？
9. a)定义一个非确定型有限状态自动机。
   b)试证：对每个非确定型有限状态自动机，存在一个确定型有限状态自动机，它们识别相同的语言。
10. a)定义一个集合 I 上的正则表达式集。
    b)解释怎么用正则表达式表示正则集合。
11. 叙述克莱因定理。
12. 试证：一个集合可由正则文法生成当且仅当它是一个正则集合。
13. 给出一个不能由有限状态自动机识别的集合的例子，并证明没有有限状态自动机能够识别它。
14. 定义一个图灵机。
15. 描述怎么用图灵机来识别集合。
16. 描述怎么用图灵机来计算数论函数。

补充练习

*1. 求识别下列每个语言的一个短语结构文法。
   a)形如 $0^{2n}1^{3n}$ 的二进制数串的集合，其中 n 是一个非负整数。
   b)二进制数串的集合：其中 0 的个数是 1 的个数的两倍。
   c)形如 w^2 的二进制数串的集合，其中 w 是二进制数串。

*2. 求产生集合$\{0^{2^n} \mid n\geq 0\}$的一个短语结构文法。

   在练习 3 和 4 中，$G=(V, T, S, P)$是一个上下文无关文法，其中 $V=\{(,), S, A, B\}$，$T=\{(,)\}$，S 是初始符号，产生式有 $S\to A$，$A\to AB$，$A\to B$，$B\to(A)$，$B\to()$，$S\to\lambda$。

3. 构造下列串的派生树。
   a)(())　　b)()(())　　c)((()()))

*4. 证明 $L(G)$就是第 3 章补充练习 57 的导语中定义的括号的合式串集合。

称一个上下文无关无法是歧义的，如果 $L(G)$中有一个词有两个派生，且将这两个派生看作带根的有序树时，产生两个不同的派生树。

5. 设文法 $G=(V, T, S, P)$为：$V=\{0, S\}$，$T=\{0\}$，S 是初始符号，产生式有 $S\to 0S$，$S\to S0$ 和 $S\to 0$。构造 0^3 的两个不同派生树，从而证明 G 是歧义的。
6. 设文法 $G=(V, T, S, P)$为：$V=\{0, S\}$，$T=(0)$，S 是初始符号，产生式有 $S\to 0S$ 和 $S\to 0$。证明 G 是非歧义的。
7. 设 A 和 B 是 V^* 的两个有限子集，其中 V 是一个字母表。问 $|AB|=|BA|$ 肯定成立吗？

8. 设 V 是一个字母表，A，B 和 C 是 V^* 的子集。证明或反驳下列各式。

a) $A(B\cup C)=AB\cup AC$　　b) $A(B\cap C)=AB\cap AC$

c) $(AB)C=A(BC)$　　d) $(A\cup B)^*=A^*\cup B^*$

9. 设 V 是一个字母表，A 和 B 是 V^* 的子集。从 $A^*\subseteq B^*$ 能否推出 $A\subseteq B$?

10. 正则表达式 $(\mathbf{2}^*)(\mathbf{0}\cup(\mathbf{12}^*))^*$ 表示的串集合是什么(串的符号在集合{0，1，2}中)？

如下递归定义集合 I 上正则表达式的*星高度* $h(\mathbf{E})$：

$h(\emptyset)=0$；

若 $\mathbf{x}\in I$，则 $\mathbf{h}(\mathbf{x})=0$；

若 $\mathbf{E}_1$，$\mathbf{E}_2$ 是正则表达式，则 $h((\mathbf{E}_1\cup\mathbf{E}_2))=h((\mathbf{E}_1\mathbf{E}_2))=\max(h(\mathbf{E}_1), h(\mathbf{E}_2))$；

若 $\mathbf{E}$ 是正则表达式，则 $h(\mathbf{E}^*)=h(\mathbf{E})+1$。

11. 求下列正则表达式的星高度：

a) $\mathbf{0}^*\mathbf{1}$　　b) $\mathbf{0}^*\mathbf{1}^*$　　c) $(\mathbf{0}^*\mathbf{01})^*$　　d) $((\mathbf{0}^*\mathbf{1})^*)^*$

e) $(\mathbf{010}^*)(\mathbf{1}^*\mathbf{01}^*)^*((\mathbf{01})^*(\mathbf{10})^*)^*$　　f) $(((((\mathbf{0}^*)\mathbf{1})^*\mathbf{0})^*)\mathbf{1})^*$

*12. 对下列每个正则表达式，求一个表示相同语言但具有最小星高度的正则表达式。

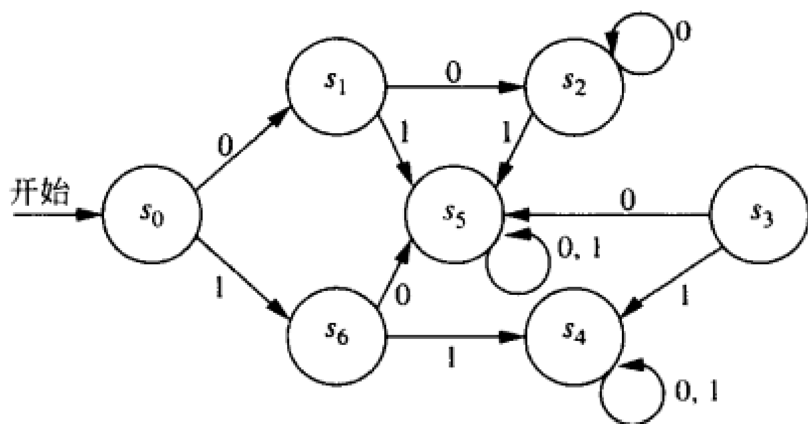
a) $(\mathbf{0}^*\mathbf{1}^*)^*$　　b) $(\mathbf{0}(\mathbf{01}^*\mathbf{0})^*)^*$　　c) $(\mathbf{0}^*\cup(\mathbf{01})^*\cup\mathbf{1}^*)^*$

13. 构造一个带输出的有限状态机，若到目前为止读到的二进制数输入串中含有 4 个或 4 个以上的 1，它则输出 1。然后再构造一个确定型有限状态自动机来识别这个集合。

14. 构造一个带输出的有限状态机器，若到目前为止读到的二进制数输入串中含有 4 个或 4 个以上连续的 1，它则输出 1。然后再构造一个确定型有限状态自动机来识别这个集合。

15. 构造一个带输出的有限状态机器，若到目前为止读到的二进制数输入串以 4 个或 4 个以上连续的 1 结尾，它则输出 1。然后再构造一个确定型有限状态自动机来识别这个集合。

16. 在有限状态机中，称状态 s' 是从状态 s *可达的*，如果存在输入串 x 使得 $f(s, x)=s'$。称状态 s 是*瞬变的*，若没有非空输入串 x 使得 $f(s, x)=s$。称状态 s 是一个*沉积点*，若对于任意输入串 x 都有 $f(s, x)=s$。对下列状态图所示的有限状态机，回答问题 a) 到 d)。



a) 哪些状态是从 s_0 可达的？

b) 哪些状态是从 s_2 可达的？

c) 哪些状态是瞬变的？

d) 哪些状态是沉积点？

*17. 设集合 S，I 和 O 都是有限集合，且 $|S|=n$，$|I|=k$，$|O|=m$。

a) 可以构造多少个不同的有限状态机(米利机) $M=(S, I, O, f, g, s_0)$(其中的初始状态 s_0 可以任意选择)？

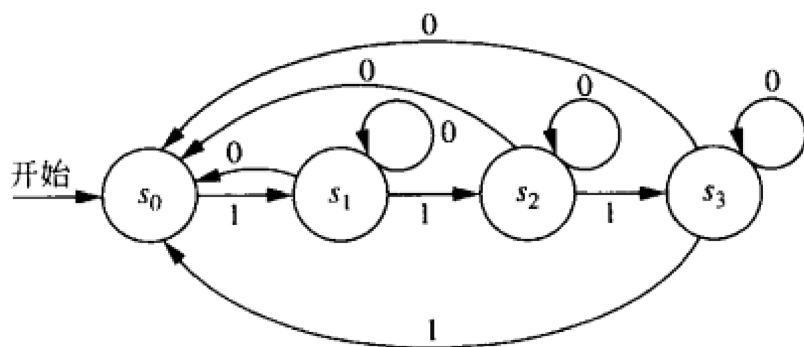
b) 可以构造多少个不同的摩尔机 $M=(S, I, O, f, g, s_0)$(其中的初始状态 s_0 可以任意选择)？

*18. 设集合 S 和 I 是有限集合，且 $|S|=n$，$|I|=k$。在下列情形下，存在多少个不同的有限状态自动机 (S, I, f, s_0, F)(其中的初始状态 s_0 以及由 S 的终结状态构成的子集 F 可以任意选择)？

a) 如果机器是确定型的。

b)如果机器是非确定型的。(注：这包括确定型机器。)

19. 对于具有如下状态图的非确定型自动机，构造一个与之等价的确定型有限状态自动机。



20. 练习 19 中的自动机识别的语言是什么？

21. 构造有限状态自动机识别下列集合：

a)$\mathbf{0^*(10)^*}$　　b)$\mathbf{(01 \cup 111)^*10^*(0 \cup 1)}$　　c)$\mathbf{(001 \cup (11)^*)^*}$

*22. 求表示由 0 和 1 组成的下列串集合的正则表达式：

a)偶数个 1 与奇数个 0 交替出现。

b)包含至少 2 个连续的 0 或 3 个连续的 1。

c)不包含 3 个连续的 0 或 2 个连续的 1。

*23. 试证：若 A 是一个正则集合，则 A 也是。

*24. 试证：若 A 和 B 都是正则集合，则 $A \cap B$ 也是。

*25. 求识别由 0 和 1 组成的下列串集合的有限状态自动机：

a)以不超过 3 个连续的 0 开始，且至少包含 2 个连续的 1。

b)包含偶数多个符号，且不含 101。

c)有 3 个由 2 个或 2 个以上的 1 组成的块，且有至少 2 个 0。

*26. 用 11.4 节的练习 16 所给的泵引理证明：$\{0^{2^n} \mid n \in \mathbf{N}\}$不是正则的。

*27. 用 11.4 节的练习 16 所给的泵引理证明：$\{1^p \mid p$ 是素数$\}$不是正则的。

*28. 对于上下文无关语言，有与正则集合的泵引理类似的结果。设 $L(G)$是上下文无关语言 G 识别的语言。此结果是：存在常量 N，如果 z 是 $L(G)$中的一个词，且 $l(w) \geqslant N$，则 z 可以写成 $uvwxy$，其中 $l(vwx) \leqslant N$，$l(vx) \geqslant 1$，且 uv^iwx^iy 属于 $L(G)(i=0, 1, 2, 3, \cdots)$。用这个结果证明：不存在上下文无关文法 $L(G)$满足$\{0^n1^n2^n \mid n=0, 1, 2, \cdots\}$。

计算机题目

写出具有下列输入和输出的程序。

1. 给定短语结构文法的产生式，根据乔姆斯基分类方法，判断此文法所在的类。

*2. 给定一个上下文无关文法的产生式和一个串，如果这个串在此文法生成的语言中，产生这个串的派生树。

3. 给定一个摩尔机的状态表和一个输入串，产生此机器生成的输出串。

4. 给定一个米利机的状态表和一个输入串，产生此机器生成的输出串。

5. 给定一个确定型有限状态自动机的状态表和一个串，判断这个串能否由此自动机识别。

6. 给定一个非确定型有限状态自动机的状态表和一个串，判断这个串能否由此自动机识别。

*7. 给定一个非确定型有限状态自动机的状态表，构造一个识别相同语言的确定型有限状态自动机的状态表。

**8. 给定一个正则表达式，构造一个非确定型有限状态自动机识别这个表达式表示的集合。

9. 给定一个正则文法，构造一个有限状态自动机识别这个文法生成的语言。

10. 给定一个有限状态自动机，构造一个正则文法生成这个自动机识别的语言。

*11. 给定一个图灵机，求一个给定的输入串所产生的输出串。

计算和研究

用一个计算程序或你已经写出的程序做下列练习。

1. 如下解两个状态的忙碌海狸问题：检查所有具有两个状态且字母表为{1，B}的图灵机。
*2. 如下解 3 个状态的忙碌海狸问题：检查所有具有 3 个状态且字母表为{1，B}的图灵机。
**3. 如下求 4 个状态的忙碌海狸机器：检查所有具有两个状态且字母表为{1，B}的图灵机。
**4. 尽力解 5 个状态的忙碌海狸问题，进展越多越好。
**5. 尽力解 6 个状态的忙碌海狸问题，进展越多越好。

写作题目

用课本以外的资料解决下列问题，并写成短文。

1. 描述怎么用一个利登梅耶系统(Lidenmeyer system)来建立某种类型植物的生长模型。利登梅耶系统用带产生式的文法来建立植物生长的各种不同方式的模型。
2. 对于各种各样的程序设计语言，如 Java，LISP，ADA 和数据库语言 SQL，给出描述其语法的巴克斯-诺尔范式规则。
3. 解释在拼写检查程序中，怎么使用有限状态机。
4. 解释在网络协议研究中，怎么使用有限状态机。
5. 解释在语音识别程序中，怎么使用有限状态机。
6. 解释概念“有限状态自动机极小化”。给出一个程序来实现这个极小化。
7. 给出细胞自动机的定义，解释它们的应用(以“生命的游戏”(Game of Life)为例)。
8. 定义下推自动机，解释怎么用下推自动机来识别集合。下推自动机能识别哪些的集合？给出验证你的答案正确性的证明概要。
9. 定义线性有界自动机，解释怎么用线性有界自动机来识别集合。线性有界自动机能识别哪些的集合？给出验证你的答案正确性的证明概要。
10. 查找图灵对现称为图灵机的机器的原始定义。他定义这样机器的动机是什么？
11. 描述“通用图灵机”(Universal Turing Machine)的概念。解释怎么构造这样的机器。
12. 解释能够用非确定型图灵机而不能用确定型图灵机的应用种类。
13. 证明一个图灵机能够模拟一个非确定型图灵机的任何动作。
14. 证明一个集合能被图灵机识别当且仅当它能由短语结构文法生成。
15. 描述 λ 演算的基本概念。解释怎么用它来研究函数的可计算性。
16. 试证：一个具有 n 个带的图灵机能做的任何事情，本章所定义的图灵机也都能做。
17. 试证：一个在两个方向都有无限带的图灵机能做的任何事情，只在一个方向有无限带的图灵机也都能做。

附录 A　指数函数和对数函数

本附录将复习指数函数和对数函数的一些基本性质。本教材从头到尾都要用到这些性质。需要对这些材料做更深入复习的学生，可以参考微积分教材或其先修教材，例如在“推荐读物”中提出的那些教材。

A.1　指数函数

设 n 是一个正整数，b 是一个固定的正实数，函数 $f_b(n)=b^n$ 是由

$$f_b(n) = b^n = b \cdot b \cdot b \cdot \cdots \cdot b$$

定义的，其中等式的右边是 n 个因子 b 相乘。

由微积分理论知：可以对任意实数 x 定义函数 $f_b(x)=b^x$。函数 $f_b(x)=b^x$ 称为以 b 为底的**指数函数**。对于以 b 为底的指数函数，当 x 不是整数时，其函数值的求法将略去。

下面的定理 1 给出了指数函数的两个重要性质。它们及其相关性质的证明可以在微积分教科书中找到。

定理 1　设 b 是一个实数，则

i) $b^{x+y}=b^x b^y$。

ii) $(b^x)^y=b^{xy}$。

图 A-1 是一些指数函数的函数图。

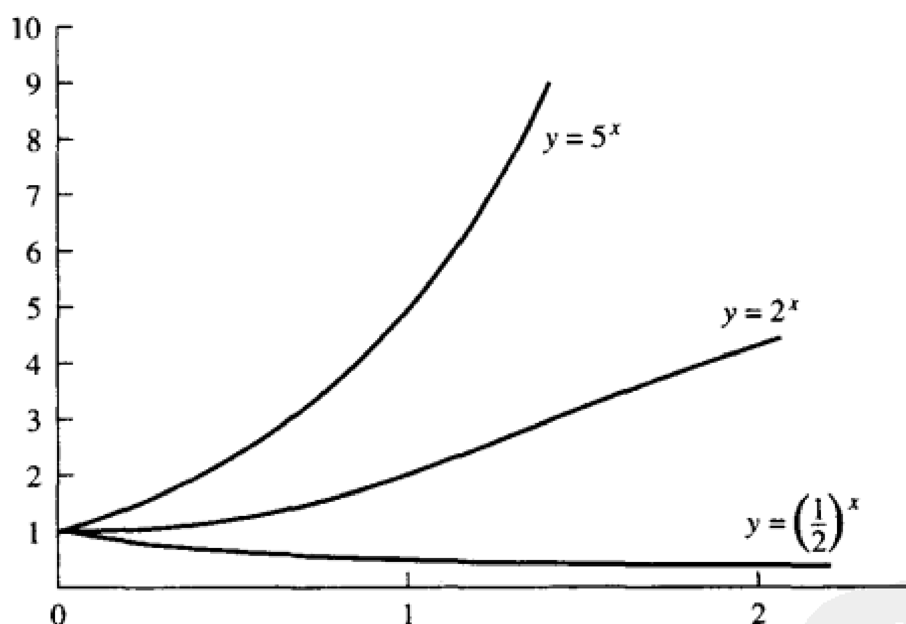


图 A-1　底为 $\frac{1}{2}$、2 和 5 的指数函数的图形

A.2　对数函数

设 b 是一个实数，且 $b>1$，则指数函数 b^x 是严格递增的（微积分证明的一个事实），且是从实数集到非负实数集的一个一一对应，因而就有反函数，此反函数称为以 b 为底的**对数函数**。换句话说，如果 b 是一个大于 1 的实数，x 是一个正实数，则

$$b^{\log_b x} = x$$

这个函数在 x 处的值称为以 b 为底 x 的**对数**。

由定义可以看出

$$\log_b b^x = x$$

定理 2 给出了对数的几个重要性质。

定理 2 设 b 是一个大于 1 的实数，则

i)当 x 和 y 都是正实数时，$\log_b(xy)=\log_b x+\log_b y$。

ii)当 x 是正实数时，$\log_b(x^y)=y\log_b x$。

证 因为 $\log_b(xy)$是满足 $b^{\log_b(xy)}=xy$ 的唯一实数，所以，为证明第 1 部分，只要证明 $b^{\log_b x+\log_b y}=xy$。根据定理 1 的第 1 部分，我们有

$$b^{\log_b x+\log_b y} = b^{\log_b x}b^{\log_b y} = xy$$

为证明第 2 部分，只要证明 $b^{y\log_b x}=x^y$。根据定理 1 的第 2 部分，我们有

$$b^{y\log_b x} = (b^{\log_b x})^y = x^y$$

□

下面定理将两个不同底的对数联系在一起。

定理 3 设 a 和 b 是大于 1 的实数，x 是正实数，则

$$\log_a x = \log_b x/\log_b a$$

证 为证明这个结果，只要证明

$$b^{\log_a x\cdot\log_b a} = x$$

根据定理 1 的第 2 部分，我们有

$$b^{\log_a x\cdot\log_b a} = (b^{\log_b a})^{\log_a x} = a^{\log_a x} = x$$

这就完成了证明。 □

本教材中，最常用的对数的底是 2，所以我们用记号 $\log x$ 来记 $\log_2 x$。

图 A-2 是函数 $f(x)=\log x$ 的图形。根据定理 3，如果底 b 不是 2，所得的函数是 $\log x$ 的常数倍，即$(1/\log b)\log x$。

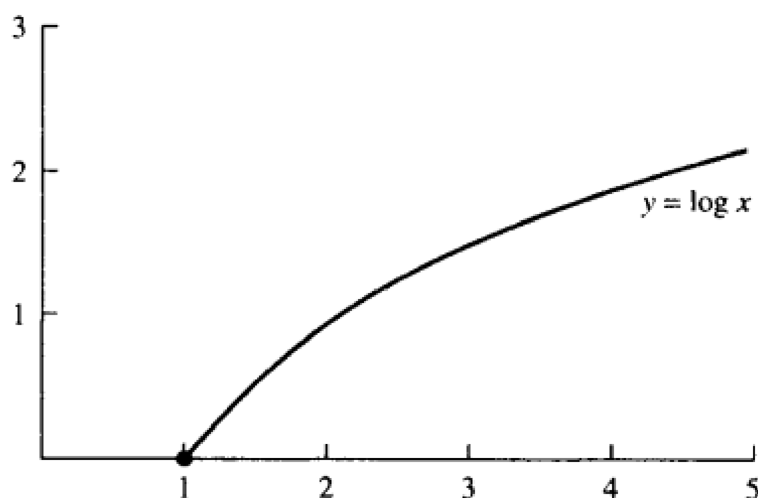


图 A-2 对数函数 $f(x)=\log x$ 的图形

练习

1. 将下列每个数表示成 2 的乘幂。

   a) $2\cdot 2^2$　　b) $(2^2)^3$　　c) $2^{(2^2)}$

2. 求下列各式的值。

   a) $\log_2 1024$　　b) $\log_2 1/4$　　c) $\log_4 8$

3. 设 $\log_4 x=y$，求下列各式的值。

   a) $\log_2 x$　　b) $\log_8 x$　　c) $\log_{16} x$

☞4. 设 a，b 和 c 是正实数。证明 $a^{\log_b c}=c^{\log_b a}$。

5. 当 b 为下列值时，画出 $f(x)=b^x$ 的图形。

   a) 3　　b) 1/3　　c) 1

6. 当 b 为下列值时，画出 $f(x)=\log_b x$ 的图形。

   a) 4　　b) 100　　c) 1000

附录B 伪 代 码



在本教材中，用英语和伪代码两种语言来描述算法。伪代码是过程步骤的英语描述和实际程序设计语言的过程说明之间的一个中间步骤。使用伪代码的优点在于它的简单性和可理解性，它很容易被写出来，也容易从它产生实际的计算机代码(用各种各样的程序设计语言)。

在这个附录中，我们描述本教材中使用的伪代码的格式和语法，这种伪代码的基本结构与Pascal的非常类似，而Pascal是目前最广泛用于教学中的一种程序设计语言。但我们使用的伪代码比正式的程序设计语言要自由得多，因为许多步骤都可以使用英语来描述。

本附录不是一个正式的学习材料，相反，应该把它当做学生的参考指南。当学生在学习课文中算法的描述时，或在写练习的伪代码求解时，可以使用它。

B.1 过程语句

算法的伪代码描述以语句 **procedure** 开始，这个语句给出算法的名称、列出输入变元和描述每个输入变元所具有的类型。例如，语句

```
procedure maximum(L: 整数列表)
```

是一个算法的伪代码描述中的第一个语句，此算法的名称是 *maximum*，它要求一列整数的最大值。

B.2 赋值语句和其他语句类型

赋值语句是用来对变元进行赋值的。在赋值语句中，左边是变元名，右边是一个表达式，表达式中可以包含常量、已被赋值的变元和过程定义的函数。右边可以包含任何常见的算术运算。本书中的伪代码还可以包含任何已经定义的运算，即使这些运算要需用实际程序设计语言中的许多语句才能实现。

赋值用符号 := 表示。这样，赋值语句的形式为

```
变元 := 表达式
```

例如，语句

```
max := a
```

将 a 的值赋给变元 *max*。也可以使用如下语句：

```
x := 列表 L 中的最大整数
```

它将 x 的值置为列表 L 中的最大整数。要将这个语句翻译成实际程序设计语言，可能需要不止一个语句。同时，可以用命令

```
交换 a 和 b
```

来交换 a 和 b。虽然能够将这个语句表示成多个赋值语句(见本附录的练习 2)，但为简单起见，我们经常使用这个伪代码的简写形式。

B.3 命令块

可以将许多语句分成块来执行复杂过程。这些块用 **begin** 和 **end** 语句来划分，块中的所有语句都按相同的间隔缩进。

```
begin
    语句 1
    语句 2
    语句 3
       ⋮
    语句 n
end
```

块中的所有语句按顺序执行。

B.4 注释

在本书的伪代码中，以花括号括起来的语句是不执行的，这样的语句被用作注释或提示，帮助解释过程是怎么运行的。例如，可以用语句

```
{x 是 L 中的最大元素}
```

提醒读者，在过程的那个地方，x 等于列表 L 中的最大元素。

B.5 条件结构

我们将使用的最简单的条件结构是

```
if 条件 then 语句
```

或者

```
if 条件 then
begin
    语句块
end
```

执行时，首先检查条件，如果它为真，则执行所给的语句。例如，在 2.1 节的算法 1(它求一个整数集合的最大值)中，用条件语句来检查 $max < a_i$ 是否对每个变元都成立，如果是，就将 a_i 的值赋给 max。

常常要用更一般的(条件)结构。这种结构不仅在一种情形下(当所给的条件为真时)能够使用，在另一种情形下(条件为假)时也能使用。这种结构是

```
if 条件 then 语句 1
else 语句 2
```

注意，语句 1 和 2 都可以用一个程序块来替代。有时候，还要用更一般形式的条件结构，这种结构是

```
if 条件 1 then 语句 1
else if 条件 2 then 语句 2
else if 条件 3 then 语句 3
    ⋮
else if 条件 n then 语句 n
else 语句 n+1
```

在使用这种结构时，若条件 1 为真，则执行语句 1，然后程序就退出此结构。如果条件 1 为假，则程序检查条件 2 是否为真；如是，则执行语句 2；等等。这样，如果前面 $n-1$ 个条件都不成立，但条件 n 成立，则执行语句 n。最后如果条件 1、条件 2、条件 3、…、条件 n 都不成立，则执行语句 $n+1$。注意，$n+1$ 个语句中的任何一个都可用一个程序块来替代。

B.6 循环结构

本书的伪代码中，有两种循环结构，第一是"for 结构"，它的形式是

```
for 变元 :=起始值  to  结束值
    语句
```

或

```
for 变元 :=起始值  to  结束值
begin
    语句块
end
```

循环开始时，若起始值小于或等于结束值，则将起始值赋给变元，并在变元的这个值下，执行结构尾部的语句。然后变元的值增加 1，并在变元的这个新值下，再执行这个语句(或语句块中的语句)。这样一直重复下去，直至变元达到结束值。在执行这个命令之后，变元等于结束值，并且算法继续进行到下一个语句。如果起始值超过结束值，循环中的任何语句都不执行。

下列伪代码用"for"循环结构求正整数 1 到 n 的和。

```
sum :=0
for i :=1 to n
    sum :=sum+1
```

本书中还用一种更一般的"for"语句，其形式是

```
for 具有某种性质的所有元素
```

它的含义是：对具有某种性质的所有元素，跟在它后面的语句或语句组将相继被执行。

我们使用的第二类循环结构是"while"结构，其形式是

```
while 条件
    语句
```

或

```
while 条件
begin
    语句块
end
```

在执行这个结构时，首先检查所给的条件，若为真，则执行其后面的语句，这次执行可能改变某些变元的值，而这些变元是条件的一部分。在这些命令执行完毕后，若条件仍然为真，则再执行这些语句。此过程一直执行到条件变成假为止。例如，可以用下列包含"while 的伪代码段来求整数 1 到 n 的和"。

```
sum := 0
while n > 0
begin
    sum := sum + n
    n := n - 1
end
```

注意，任何"for"结构都可转换为"while"结构(见本附录结尾的练习 3)，但"for"结构更容易理解。所以，只要适合时，就优先使用"for"结构而不是对应的"while"结构。

B.7　在过程中使用其他过程

在一个过程内部还可以使用其他过程(或在递归程序中使用其自身)，方法是写出这个过程的名称，后面再跟上此过程的输入。例如：

```
max(L)
```

将执行输入为 L 的过程 max。当这个过程中的所有步骤都执行完毕后，再继续执行这个过程的下一个语句。

练习

1. 下列两个赋值语句块的区别是什么？

   $a := b$
   $b := c$

   和

   $b := c$
   $a := b$

2. 用赋值语句构造一个过程来交换变元 x 和 y 的值。所需赋值语句的最少个数是什么？
3. 说明怎样用"while"结构来写形式为

   for $i :=$ 起始值 **to** 结束值
   　　语句

   的循环？

奇数练习题答案

第1章

1.1节

1. a)是，T　b)是，F　c)是，T　d)是，F
   e)不是　f)不是　g)是，T
3. a)今天不是星期四。
   b)新泽西州有污染。
   c)$2+1\neq 3$。
   d)缅因州的夏天不热或阳光不明媚。
5. a)在海岸附近没发现过鲨鱼。
   b)在新泽西海岸游泳是允许的，并且在海岸附近发现过鲨鱼。
   c)在新泽西海岸不允许游泳，或者在海岸附近发现过鲨鱼。
   d)如果在新泽西海岸游泳是允许的，那么在海岸附近没发现过鲨鱼。
   e)如果在海岸附近没发现过鲨鱼，那么在新泽西海岸游泳是允许的。
   f)如果在新泽西海岸不允许游泳，那么在海岸附近没发现过鲨鱼。
   g)在新泽西海岸允许游泳当且仅当在海岸附近没发现过鲨鱼。
   h)在新泽西海岸不允许游泳，并且在新泽西海岸允许游泳或者在海岸附近没发现过鲨鱼。
7. a)$p\wedge q$　b)$p\wedge\neg q$　c)$\neg p\wedge\neg q$
   d)$p\vee q$　e)$p\rightarrow q$　f)$(p\vee q)\wedge(p\rightarrow\neg q)$　g)$q\leftrightarrow p$
9. a)$\neg p$　b)$p\wedge\neg q$　c)$p\rightarrow q$　d)$\neg p\rightarrow\neg q$
   e)$p\rightarrow q$　f)$q\wedge\neg p$　g)$q\rightarrow p$
11. a)$r\wedge\neg p$　b)$\neg p\wedge q\wedge r$　c)$r\rightarrow(q\leftrightarrow\neg p)$
    d)$\neg q\wedge\neg p\wedge r$　e)$(q\rightarrow(\neg r\wedge\neg p))\wedge\neg((\neg r\wedge\neg p)\rightarrow q)$　f)$(p\wedge r)\rightarrow\neg q$
13. a)False(假)　b)True(真)　c)True(真)　d)True(真)
    e)True(真)　f)True(真)　g)False(假)　h)True(真)
15. a)同或：如果你学过微积分或计算机科学，或两者都学过，就可以选修离散数学。异或：如果你学过微积分或计算机科学，但并非两者都学过，就可以选修离散数学。这里很可能想表示的是同或。
    b)同或：你可以拿回扣，也可以得低息贷款，你也可以既拿回扣又得低息贷款。异或：你可以拿回扣，也可以得低息贷款，但不能既拿回扣，又得低息贷款。这里很可能想表示的是异或。
    c)同或：你可以从A列选两项，不选B列；你也可以不选A列而从B列选三项；你还可以共选五项，包括A列两项和B列三项。异或：你可以从A列选两项，也可以从B列选三项，但不能都选。这里很可能想表示的是异或。
    d)同或：2英尺多的雪或冷风零下100度(华氏)，或两者均具备，学校将停课。异或：2英尺多的雪或冷风零下100度，但并非两者均具备，学校将停课。这里很可能想表示的是同或。
17. a)东北风吹就下雪。
    b)温暖若能持续一周，苹果树就会开花。
    c)若活塞队赢得冠军，那么他们打败了湖人队。
    d)如果你登上了朗峰顶，那你必定已走了8英里(1英里=1.6公里)。

e)如果你世界闻名，就能做终身教授。

f)如果你驾车超过 400 英里，就得买汽油了。

g)如果你的保修单有效，你购买 CD 机必定还不足 90 天。

19. a)当且仅当外边很热时你才能买冰激淋卷。

b)当且仅当你手持唯一的胜券时才能赢得比赛。

c)当且仅当你有关系时才会得到提拔。

d)当且仅当你看电视时思想才会衰退。

e)当且仅当我乘坐火车的日子，它才会晚点。

21. a)逆命题："只有今天下雪，我明天才去滑雪。"逆否命题："如果我明天不滑雪，那么今天一定没有下雪。"否命题："如果今天不下雪，那么我明天不滑雪。"

b)逆命题："如果我今天来上课，就会有测验。"逆否命题："如果我今天不来上课，就不会有测验。"否命题："如果不会有测验，我就不来上课。"

c)逆命题："如果一个正整数没有 1 和它自身以外的因数，它就是素数。"逆否命题："如果一个正整数有既不是 1 又不是它自己的因数，它不是素数。"否命题："如果一个正整数不是素数，那它有 1 和它自身以外的因数。"

23.

a)

p	$\neg p$	$p \wedge \neg p$
T	F	F
F	T	F

b)

p	$\neg p$	$p \vee \neg p$
T	F	T
F	T	T

c)

p	q	$\neg q$	$p \vee \neg q$	$(p \vee \neg q) \to q$
T	T	F	T	T
T	F	T	T	F
F	T	F	F	T
F	F	T	T	F

d)

p	q	$p \vee q$	$p \wedge q$	$(p \vee q) \to (p \wedge q)$
T	T	T	T	T
T	F	T	F	F
F	T	T	F	F
F	F	F	F	T

e)

p	q	$p \to q$	$\neg q$	$\neg p$	$\neg q \to \neg p$	$(p \to q) \leftrightarrow (\neg q \to \neg p)$
T	T	T	F	F	T	T
T	F	F	T	F	F	T
F	T	T	F	T	T	T
F	F	T	T	T	T	T

f)

p	q	$p \to q$	$q \to p$	$(p \to q) \to (q \to p)$
T	T	T	T	T
T	F	F	T	T
F	T	T	F	F
F	F	T	T	T

25. 对 a)、b)、c)、d)和 f)有如下表：

p	q	$(p\vee q)\to(p\oplus q)$	$(p\oplus q)\to(p\wedge q)$	$(p\vee q)\oplus(p\wedge q)$	$(p\leftrightarrow q)\oplus(\neg p\leftrightarrow q)$	$(p\oplus q)\to(p\oplus\neg q)$
T	T	F	T	F	T	T
T	F	T	F	T	T	F
F	T	T	F	T	T	F
F	F	T	T	F	T	T

对 e)有下表：

p	q	r	$\neg p$	$\neg r$	$p\leftrightarrow q$	$\neg p\leftrightarrow\neg r$	$(p\leftrightarrow q)\oplus(\neg p\leftrightarrow\neg r)$
T	T	T	F	F	T	T	F
T	T	F	F	T	T	F	T
T	F	T	F	F	F	T	T
T	F	F	F	T	F	F	F
F	T	T	T	F	F	F	F
F	T	F	T	T	F	T	T
F	F	T	T	F	T	F	T
F	F	F	T	T	T	T	F

27.

p	q	$p\to\neg q$	$\neg p\leftrightarrow q$	$(p\to q)\vee(\neg p\to q)$	$(p\to q)\wedge(\neg p\to q)$	$(p\leftrightarrow q)\vee(\neg p\leftrightarrow q)$	$(\neg p\leftrightarrow\neg q)\leftrightarrow(p\leftrightarrow q)$
T	T	F	F	T	T	T	T
T	F	T	T	T	F	T	T
F	T	T	T	T	T	T	T
F	F	T	F	T	F	T	T

29.

p	q	r	$p\to(\neg q\vee r)$	$\neg p\to(q\to r)$	$(p\to q)\vee(\neg p\to r)$	$(p\to q)\wedge(\neg p\to r)$	$(p\leftrightarrow q)\vee(\neg q\leftrightarrow r)$	$(\neg p\leftrightarrow\neg q)\leftrightarrow(q\leftrightarrow r)$
T	T	T	T	T	T	T	T	T
T	T	F	F	T	T	T	T	F
T	F	T	T	T	T	F	T	T
T	F	F	T	T	T	F	F	F
F	T	T	T	T	T	T	F	F
F	T	F	T	F	T	F	T	T
F	F	T	T	T	T	T	T	F
F	F	F	T	T	T	F	T	T

31.

p	q	r	s	$p\leftrightarrow q$	$r\leftrightarrow s$	$(p\leftrightarrow q)\leftrightarrow(r\leftrightarrow s)$
T	T	T	T	T	T	T
T	T	T	F	T	F	F
T	T	F	T	T	F	F
T	T	F	F	T	T	T
T	F	T	T	F	T	F
T	F	T	F	F	F	T
T	F	F	T	F	F	T
T	F	F	F	F	T	F
F	T	T	T	F	T	F

（续）

p	q	r	s	$p\leftrightarrow q$	$r\leftrightarrow s$	$(p\leftrightarrow q)\leftrightarrow(r\leftrightarrow s)$
F	T	T	F	F	F	T
F	T	F	T	F	F	T
F	T	F	F	F	T	F
F	F	T	T	T	T	T
F	F	T	F	T	F	F
F	F	F	T	T	F	F
F	F	F	F	T	T	T

33. a)按位或是 111 1111；按位并是 000 0000；按位异或(按位加)是 111 1111。
    b)按位或是 1111 1010；按位并是 1010 0000；按位异或(按位加)是 0101 1010。
    c)按位或是 10 0111 1001；按位并是 00 0100 0000；按位异或(按位加)是 10 0011 1001。
    d)按位或是 11 1111 1111 ；按位并是 00 0000 0000；按位异或(按位加)是 11 1111 1111。
35. 0.2，0.6
37. 0.8，0.6
39. a)第 99 条语句为真，其余语句为假。
    b)语句 1 至 50 均为真，语句 51 至 99 均为假。
    c)这不可能发生，这是个悖论。
41. “如果我问你右边的路是否通向废墟，你会说是吗？”
43. a) $q\rightarrow p$　　b) $q\wedge\neg p$　　c) $q\rightarrow p$　　d) $\neg q\rightarrow\neg p$
45. 不一致。
47. 一致
49. NEW **AND** JERSEY **AND** BEACHES，(JERSEY **AND** BEACHES) **NOT** NEW
51. A 是武士，B 是流氓。
53. A 是武士且 B 是武士。
55. A 是流氓，B 是武士。
57. 按薪水减少的顺序：傅雷德、麦吉、杰尼斯。
59. 侦探可以断定男管家和厨师说谎，但不能判断究竟是园丁还是杂役说真话。
61. 日本人养斑马，挪威人喝矿泉水。

1.2 节

1. 等价关系由下表中相应两列的一致推出。

p	$p\wedge T$	$p\vee F$	$p\wedge F$	$p\vee T$	$p\vee p$	$p\wedge p$
T	T	T	F	T	T	T
F	F	F	F	T	F	F

3. a)

p	q	$p\vee q$	$q\vee p$
T	T	T	T
T	F	T	T
F	T	T	T
F	F	F	F

b)

p	q	$p\wedge q$	$q\wedge p$
T	T	T	T
T	F	F	F
F	T	F	F
F	F	F	F

5.

p	q	r	$q\vee r$	$p\wedge(q\vee r)$	$p\wedge q$	$p\wedge r$	$(p\wedge q)\vee(p\wedge r)$
T	T	T	T	T	T	T	T
T	T	F	T	T	T	F	T
T	F	T	T	T	F	T	T
T	F	F	F	F	F	F	F
F	T	T	T	F	F	F	F
F	T	F	T	F	F	F	F
F	F	T	T	F	F	F	F
F	F	F	F	F	F	F	F

7. a)

p	q	$p\wedge q$	$(p\wedge q)\to p$
T	T	T	T
T	F	F	T
F	T	F	T
F	F	F	T

b)

p	q	$p\vee q$	$p\to(p\vee q)$
T	T	T	T
T	F	T	T
F	T	T	T
F	F	F	T

c)

p	q	$\neg p$	$p\to q$	$\neg p\to(p\to q)$
T	T	F	T	T
T	F	F	F	T
F	T	T	T	T
F	F	T	T	T

d)

p	q	$p\wedge q$	$p\to q$	$(p\wedge q)\to(p\to q)$
T	T	T	T	T
T	F	F	F	T
F	T	F	T	T
F	F	F	T	T

e)

p	q	$p\to q$	$\neg(p\to q)$	$\neg(p\to q)\to p$
T	T	T	F	T
T	F	F	T	T
F	T	T	F	T
F	F	T	F	T

f)

p	q	$p\to q$	$\neg(p\to q)$	$\neg q$	$\neg(p\to q)\to\neg q$
T	T	T	F	F	T
T	F	F	T	T	T
F	T	T	F	F	T
F	F	T	F	T	T

9. 对每种情况都证明只要前提成立，结论就成立。

a)如果前题 $p\wedge q$ 成真，那么根据合取的定义，结论 p 必为真。

b)如果前题 p 成真，根据析取的定义，结论 $p\vee q$ 也为真。

c)如果前题 $\neg p$ 成真，亦即 p 为假，那么结论 $p\to q$ 为真。

d)如果前题 $p\wedge q$ 为真，那么 p 和 q 均为真，所以结论 $p\to q$ 为真。

e)如果前题 $\neg(p\to q)$为真，那么 $p\to q$ 为假，所以 p 为真(同时 q 为假)。

f)如果前题 $\neg(p\to q)$为真，那么 $p\to q$ 为假，所以 p 为真而 q 为假。因而结论 $\neg q$ 为真。

11. a)如果 p 为真，那么 $p\vee(p\wedge q)$为真，因为析取中第一项为真。另一方面，如果 p 为假，那么 $p\wedge q$ 为假，所以 $p\vee(p\wedge q)$也为假。由于 p 和 $p\vee(p\wedge q)$总是有同样的真值，它们等价。

b)如果 p 为假，那么 $p\wedge(p\vee q)$也为假，因为合取中第一项为假。另一方面，如果 p 为真，那么由于 $p\vee q$ 为真，所以合取的两项均为真。由于 p 和 $p\wedge(p\vee q)$总是有同样的真值，它们等价。

p	q	$p\wedge q$	$p\vee(p\wedge q)$	$p\vee q$	$p\wedge(p\vee q)$
T	T	T	T	T	T
T	F	F	T	T	T
F	T	F	F	T	F
F	F	F	F	F	F

13. 唯一能使这一蕴含关系为假的情况是 $\neg q\wedge(p\to q)$ 为真而 $\neg p$ 为假。由 $\neg p$ 为假知 p 必为真。由 $\neg q\wedge(p\to q)$ 为真，知 $\neg q$ 必为真，从而 q 为假。由 p 为真知 $p\to q$ 必为假，而这是不可能的。
15. 这些命题不可能逻辑等价，因为当 p、q 和 r 全为假时，$(p\to q)\to r$ 也为假，但 $p\to(q\to r)$ 为真。
17. 命题 $\neg p\leftrightarrow q$ 为真当且仅当 $\neg p$ 和 q 有同样的真值，也就是 p 和 q 有不同的真值。类似地，$p\leftrightarrow\neg q$ 在完全同样的情况下为真。所以这两个表达式逻辑等价。
19. 命题 $\neg(p\leftrightarrow q)$ 为真当且仅当 $p\leftrightarrow q$ 为假，即 p 和 q 有不同的真值。由于这恰恰就是 $\neg p\leftrightarrow q$ 为真的情况，所以这两个表达式逻辑等价。
21. 为了使 $(p\to r)\wedge(q\to r)$ 为假，两个蕴含式之一必为假，当 r 为假且 p 和 q 至少有一个为真时恰会发生这种情况。但这种情形下 $p\vee q$ 为真且 r 为假，当 $(p\vee q)\to r$ 为假时满足条件。由于两个命题恰在相同情形下为假，因此它们逻辑等价。
23. 为了使 $(p\to r)\vee(q\to r)$ 为假，两个蕴含式都必须为假，当 r 为假且 p 和 q 都为真时恰会发生这种情况。但这种情形下 $p\wedge q$ 为真且 r 为假，当 $(p\wedge q)\to r$ 为假时满足条件。由于两个命题恰在相同情形下为假，因此它们逻辑等价。
25. 这一事实在 1.1 节中定义双蕴含时已看到过。当 p 和 q 有相同的真值时两边都为真。
27. 当 p 和 q 有相反的真值时两边恰好都为真。
29. 见下表，最后一列全为 T。

p	q	r	$p\to q$	$q\to r$	$(p\to q)\wedge(q\to r)$	$p\to r$	$(p\to q)\wedge(q\to r)\to(p\to r)$
T	T	T	T	T	T	T	T
T	T	F	T	F	F	F	T
T	F	T	F	T	F	T	T
T	F	F	F	T	F	F	T
F	T	T	T	T	T	T	T
F	T	F	T	F	F	T	T
F	F	T	T	T	T	T	T
F	F	F	T	T	T	T	T

31. 如果取两次对偶，那么每个 $\vee$ 先变成 $\wedge$，又变回 $\vee$；每个 $\wedge$ 先变成 $\vee$，又变回 $\wedge$；每个 **T** 先变为 **F**，再变回 **T**；每个 **F** 先变为 **T**，再变回 **F**。因此 $(s^*)^*=s$。
33. 令 p 和 q 为只含运算符 $\wedge$、$\vee$ 和 $\neg$，以及 **T** 和 **F** 的等价复合命题，注意 $\neg p$ 和 $\neg q$ 也等价。反复用德摩根定律尽可能把这些复合命题中推进，同时改 $\vee$ 为 $\wedge$，改 $\wedge$ 为 $\vee$，改 **T** 为 **F**，改 **F** 为 **T**。这样能证明 $\neg p$ 和 $\neg q$ 与 p^* 和 q^* 是一样的，只是其中的原子命题被它的非命题所代替。由此可断定 p^* 和 q^* 也等价，因为 $\neg p$ 和 $\neg q$ 等价。
35. $(p\wedge q\wedge\neg r)\vee(p\wedge\neg q\wedge r)\vee(\neg p\wedge q\wedge r)$
37. 对给定的复合命题 p，构造其真值表，然后用析取范式写下一个逻辑等价于 p 的命题 q。由于 q 只涉及 $\neg$、$\wedge$ 和 $\vee$，这就证明了这三个运算符构成一个功能完全集。
39. 根据练习 27，给定一个复合命题 p，我们可以写出一个只含 $\neg$、$\wedge$ 和 $\vee$，且与 p 逻辑等价的命题 q。根据德摩根定律，我们可以用 $\neg(\neg p_1\vee\neg p_2\vee\cdots\vee\neg p_n)$ 取代 $(p_1\wedge p_2\wedge\cdots\wedge p_n)$，从而消去所有 $\wedge$。
41. 在 p 或 q 或两者均为假时，$\neg(p\wedge q)$ 为真；当 p 和 q 均为真时，$\neg(p\wedge q)$ 为假。由于这就是 $p\mid q$ 的定义，所以这两个复合命题逻辑等价。
43. $\neg(p\vee q)$ 在 p 和 q 均为假时为真，否则为假。由于这是 $p\downarrow q$ 的定义，两者逻辑等价。
45. $((p\downarrow p)\downarrow q)\downarrow((p\downarrow p)\downarrow q)$
47. 从真值表或 $p\mid q$ 的定义可直接得出结论。
49. 16
51. 若数据库是打开的，那么或者系统置于初始状态，或者监督程序置于关闭状态。
53. 全部 9 个。
55. 要判断 c 是否为永真式，将判断可满足性的算法应用于 $\neg c$。如果算法表明 $\neg c$ 是可满足的，则可知 c 不是永真式；如果算法称 $\neg c$ 不是可满足的，则 c 是永真式。

1.3节

1. a)T b)T c)F

3. a)T b)F c)F d)F

5. a)有个学生除周末外每天都花5个多小时在课堂上。
   b)所有学生除周末外每天都花5个多小时在课堂上。
   c)有个学生并非除周末外每天都花5个多小时在课堂上。
   d)没有学生除周末外每天都花5个多小时在课堂上。

7. a)每个喜剧演员都很有趣。
   b)每个人都是有趣的喜剧演员。
   c)存在某个人，若她或他是喜剧演员，那么她或他很有趣。
   d)一些喜剧演员很有趣。

9. a) $\exists x(P(x) \land Q(x))$ b) $\exists x(P(x) \land \neg Q(x))$ c) $\forall x(P(x) \lor Q(x))$ d) $\forall x \neg(P(x) \lor Q(x))$

11. a)T b)T c)F d)F e)T f)F

13. a)True b)True c)True d)True

15. a)True b)False c)True d)False

17. a) $P(0) \lor P(1) \lor P(2) \lor P(3) \lor P(4)$ b) $P(0) \land P(1) \land P(2) \land P(3) \land P(4)$
    c) $\neg P(0) \lor \neg P(1) \lor \neg P(2) \lor \neg P(3) \lor \neg P(4)$ d) $\neg P(0) \land \neg P(1) \land \neg P(2) \land \neg P(3) \land \neg P(4)$
    e) $\neg(P(0) \lor P(1) \lor P(2) \lor P(3) \lor P(4))$ f) $\neg(P(0) \land P(1) \land P(2) \land P(3) \land P(4))$

19. a) $P(1) \lor P(2) \lor P(3) \lor P(4) \lor P(5)$ b) $P(1) \land P(2) \land P(3) \land P(4) \land P(5)$
    c) $\neg(P(1) \lor P(2) \lor P(3) \lor P(4) \lor P(5))$ d) $\neg(P(1) \land P(2) \land P(3) \land P(4) \land P(5))$
    e) $(P(1) \land P(2) \land P(4) \land P(5)) \lor (\neg P(1) \lor \neg P(2) \lor \neg P(3) \lor \neg P(4) \lor \neg P(5))$

21. 令 $C(x)$ 为命题函数“x 在你的班上”。
    a) $\exists x H(x)$ 和 $\exists x(C(x) \land H(x))$，其中 $H(x)$ 是“x 会说印地语”。
    b) $\forall x F(x)$ 和 $\forall x(C(x) \to F(x))$，其中 $F(x)$ 是“x 很友好”。
    c) $\exists x \neg B(x)$ 和 $\exists x(C(x) \land \neg B(x))$。其中 $B(x)$ 是“x 出生在加利福尼亚”。
    d) $\exists x M(x)$ 和 $\exists x(C(x) \land M(x))$，其中 $M(x)$ 是“x 曾演过电影”。
    e) $\forall x \neg L(x)$ 和 $\forall x(C(x) \to \neg L(x))$，其中 $L(x)$ 是“x 上过逻辑编程课程”。

23. 令 $P(x)$ 为“x 是完美的”；$F(x)$ 为“x 是你的朋友”；论域为所有人。
    a) $\forall x \neg P(x)$ b) $\neg \forall x P(x)$ c) $\forall x(F(x) \to P(x))$
    d) $\exists x(F(x) \land P(x))$，假定这意味着“至少有一个”你的朋友是完美的。
    e) $\forall x(F(x) \land P(x))$ 或 $(\forall x F(x)) \land (\forall x P(x))$
    f) $(\forall x \neg F(x)) \lor (\exists x \neg P(x))$

25. 令 $Y(x)$ 为命题函数，x 在你的学校或班上。
    a)如果我们令 $V(x)$ 为“x 住在越南”，则我们有 $\exists x V(x)$（如果论域只是你的校友），或者是 $\exists x(Y(x) \land V(x))$（如果论域是所有人）。如果令 $D(x, y)$ 表示某人 x 住在国家 y，则可以重写后者为 $\exists x(Y(x) \land D(x, 越南))$。
    b)如果令 $H(x)$ 为“x 会说印地语”，则我们有 $\exists x \neg H(x)$（论域是你的校友），或者是 $\exists x(Y(x) \land \neg H(x))$（论域是所有人）。如果令 $S(x, y)$ 表示某人 x 会说语言 y，则可以重写后者为 $\exists x(Y(x) \land \neg S(x, 印地语))$。
    c)如果令 $J(x)$、$P(x)$ 和 $C(x)$ 为命题函数，分别表示 x 会 Java、Prolog 和 C++。则有 $\exists x(J(x) \land P(x) \land C(x))$（如果论域是你校友），或者有 $\exists x(Y(x) \land J(x) \land P(x) \land C(x))$（如果论域是所有人）。如果令 $K(x, y)$ 表示某人 x 会编程语言 y，则可以重写后者为 $\exists x(Y(x) \land K(x, \text{Java}) \land K(x, \text{Prolog}) \land$

$K(x, \text{C++}))$。

d)如果令 $T(x)$为"x喜欢泰国食物"，则我们有 $\forall xT(x)$(如果论域是你的同班同学)，或者有 $\forall x(Y(x) \rightarrow T(x))$(如果论域是所有人)。如果令 $\exists(x, y)$表示某人 x 喜欢类型 y 的食物，则可以重写后者为 $\forall x(Y(x) \rightarrow E(x, \text{泰国}))$。

e)如果令 $H(x)$为"x打曲棍球"，则我们有 $\exists x \neg H(x)$(如果论域是你的同班同学)，或者有 $\exists x(Y(x) \wedge \neg H(x))$(如果论域是所有人)。如果令 $P(x, y)$表示某人 x 玩游戏 y，则可以重写后者为 $\exists x(Y(x) \wedge \neg P(x, \text{曲棍球}))$。

27. 令 $T(x)$表示 x 是永真式，$C(x)$表示 x 是矛盾。

a) $\exists xT(x)$

b) $\forall x(C(x) \rightarrow T(\neg x))$

c) $\exists x \exists y(\neg T(x) \wedge \neg C(x) \wedge \neg T(y) \wedge \neg C(y) \wedge T(x \vee y))$

d) $\forall x \forall y((T(x) \wedge T(y)) \rightarrow T(x \wedge y))$

29. a) $Q(0, 0, 0) \wedge Q(0, 1, 0)$ b) $Q(0, 1, 1) \vee Q(1, 1, 1) \vee Q(2, 1, 1)$

c) $\neg Q(0, 0, 0) \vee \neg Q(0, 0, 1)$ d) $\neg Q(0, 0, 1) \vee \neg Q(1, 0, 1) \vee \neg Q(2, 0, 1)$

31. a)令 $T(x)$为谓词表示 x 会学习新技巧，论域为年老的狗。原命题为 $\exists xT(x)$，否定为 $\forall x \neg T(x)$："没有年老的狗会学习新的技巧"。

b)令 $C(x)$为谓词，表示 x 会微积分，论域为兔子。原命题为 $\neg \exists xC(x)$，否定为 $\exists xC(x)$："有一只兔子会微积分"。

c)令 $F(x)$为谓词，表示 x 会飞，论域为鸟类。原命题为 $\forall xF(x)$，否定为 $\exists x \neg F(x)$："有一只鸟不会飞"。

d)令 $T(x)$为谓词，表示 x 会说话，论域为狗。原命题为 $\neg \exists xT(x)$，否定为 $\exists xT(x)$："有一只狗会说话"。

e)令 $F(x)$和 $R(x)$为谓词，分别表示 x 会法语和俄语，论域为班上的人。原命题为 $\neg \exists x(F(x) \wedge R(x))$，否定为 $\exists x(F(x) \wedge R(x))$："班上有人会法语和俄语"。

33. a)没有反例 b) $x=0$ c) $x=2$

35. a) $\forall x((F(x, 25\,000) \vee S(x, 25)) \rightarrow E(x))$，其中 $E(x)$是"某人 x 在特定的年份中是高贵乘客"，$F(x, y)$是"x在特定的年份飞行里程超过 y 英里"，$S(x, y)$是"x在特定年份坐飞机超过 y 次"。

b) $\forall x(((M(x) \wedge T(x, 3)) \vee (\neg M(x) \wedge T(x, 3.5))) \rightarrow Q(x))$，其中 $Q(x)$是"某人 x 有资格参加马拉松"，$M(x)$是"x是男人"，$T(x, y)$是"x跑马拉松的时间不超过 y 小时"。

c) $M \rightarrow ((H(60) \vee (H(45) \wedge T)) \wedge \forall yG(B, y))$，其中 M 是命题"学生取得硕士学位"，$H(x)$是"学生至少修过 x 个学分"，T 是命题"学生写了论文"，$G(x, y)$是"学生在课程 y 上的成绩为 x 或更高"。

d) $\exists x((T(x, 21) \wedge G(x, 4.0))$，其中 $T(x, y)$是"某人 x 修了多于 y 个学分"，$G(x, p)$是"x获得了平均为 p 的成绩"。

37. a)如果某台打印机不能提供服务且很忙，则一些作业丢失了。

b)如果每台打印机都很忙，则有作业在队列中。

c)如果有作业在队列中且丢失了，则一些打印机不能提供服务。

d)如果每台打印机都很忙且每个作业都在队列中，则一些作业丢失了。

39. a) $(\exists xF(x, 10)) \rightarrow \exists xS(x)$，其中 $F(x, y)$是"磁盘 x 有多于 ykB的空闲空间"，$S(x)$是"邮件消息 x 可以被保存"。

b) $(\exists xA(x)) \rightarrow \forall x(Q(x) \rightarrow T(x))$，其中 $A(x)$是"报警 x 是主动的"，$Q(x)$是"消息 x 被排队"，$T(x)$是"消息 x 被传送"。

c) $\forall x((x \neq \text{主控制台}) \rightarrow T(x))$，其中 $T(x)$是"诊断监控器跟踪系统 x 的状态"。

d) $\forall x(\neg L(x) \rightarrow B(x))$，其中 $L(x)$是"电话会议的主叫方将参与者 x 放到特殊列表中"，$B(x)$是"参与者 x 要支付费用"。

41. 并不总是如此。

43. x 论域中至少有一个值使 $P(x)$ 和 $Q(x)$ 中至少一个成真时，两个语句均为真。

45. a)若 A 为真，两边都逻辑等价于 $\forall xP(x)$。若 A 为假，左边显然为假。此外，对每个 x，$P(x)\land A$ 为假，所以右边为假。因此两边逻辑等价。

    b)若 A 为真，两边均逻辑等价于 $\exists xP(x)$。若 A 为假，左边显然为假。此外，对每个 x，$P(x)\land A$ 为假。所以 $\exists x(P(x)\land A)$ 为假。因此两边逻辑等价。

47. 为证明不是逻辑等价，令 $P(x)$ 为语句"x 为正"，$Q(x)$ 为语句"x 为负"，论域为整数集合。于是 $\exists xP(x)\land\exists xQ(x)$ 为真，但 $\exists x(P(x)\land Q(x))$ 为假。

49. a)True

    b)False，除非论域只含一个元素。

    c)True

51. a)Yes　　b)No　　c)juana，kiko

    d)math273，cs301　　e)juana，kiko

53. sibling(X, Y) :-mother(M, X), mother(M, Y), father(F, X), father(F, Y)

55. a) $\forall x(P(x)\to\neg Q(x))$　　b) $\forall x(Q(x)\to R(x))$　　c) $\forall x(P(x)\to\neg R(x))$

    d)不能推出。有可能有爱虚荣的教授，因为前提并没有排除无知者以外爱虚荣的人。

57. a) $\forall x(P(x)\to\neg Q(x))$　　b) $\forall x(R(x)\to\neg S(x))$　　c) $\forall x(\neg Q(x)\to S(x))$

    d) $\forall x(P(x)\to\neg R(x))$

    e)可以推出。假设 x 是婴儿，则由第一个前提，x 行为不合逻辑，因此由第三个前提，x 被人轻视。第二个前提说如果 x 能管理鳄鱼，x 就不会被人轻视。因此，x 不能管理鳄鱼。

1.4 节

1. a)对每个实数 x，存在实数 y 使得 x 小于 y。

   b)对每个实数 x 和实数 y，如果 x 和 y 都是非负的，则它们的积是非负的。

   c)对每个实数 x 和实数 y，存在实数 z 使得 $xy=z$。

3. a)班上的某个学生已经给班上的另一个学生发送了消息。

   b)班上的某个学生已经给班上的每个学生都发送了消息。

   c)班上的每个学生已经给班上的至少一个学生发送了消息。

   d)班上有某个学生已经被班上的每个学生都发送了消息。

   e)班上的每个学生已经被班上的至少一个学生发送了消息。

   f)班上的每个学生已经给班上的每个学生都发送了消息。

5. a)Sarah Smith 已经访问过 www.att.com。

   b)至少有一个人已经访问过 www.imdb.org。

   c)Jose Orez 已经访问了至少一个网站。

   d)至少有一个网站 Ashok Puri 和 Cindy Yoon 都访问过了。

   e)除了 David Belcher，还有一个人访问过 David Belcher 已经访问过的所有网站。

   f)有两个不同的人已经访问过完全相同的网站。

7. a)Abdallah Hussein 不喜欢日本菜。

   b)学校的某些学生喜欢韩国菜，学校的每个人都喜欢墨西哥菜。

   c)有一些菜肴不是 Monique Arsenault 喜欢就是 Jay Johnson 喜欢。

   d)对学校中任意一对不同的学生，他们中至少有一个不喜欢某种菜。

   e)学校中有两个学生喜欢恰恰相同的菜肴集合。

   f)对学校中的每对学生，他们对某种菜肴有相同的看法(或者都喜欢，或者都不喜欢)。

9. a) $\forall xL(x, \text{Jerry})$ b) $\forall x\exists yL(x, y)$ c) $\exists y\forall xL(x, y)$
d) $\forall x\exists y\neg L(x, y)$ e) $\exists x\neg L(\text{Lydia}, x)$ f) $\exists x\forall y\neg L(y, x)$
g) $\exists x(\forall yL(y, x)\wedge\forall z((\forall wL(w, z))\rightarrow z=x))$
h) $\exists x\exists y(x\neq y\wedge L(\text{Lynn}, x)\wedge L(\text{Lynn}, y)\wedge\forall z(L(\text{Lynn}, z)\rightarrow(z=x\vee z=y)))$
i) $\forall xL(x, x)$ j) $\exists x\forall y(L(x, y)\leftrightarrow x=y)$

11. a) $A(\text{Lois}, \text{Professor Michaels})$ b) $\forall x(S(x)\rightarrow A(x, \text{Professor Gross}))$
c) $\forall x(F(x)\rightarrow(A(x, \text{Professor Miller})\vee A(\text{Professor Miller}, x)))$
d) $\exists x(S(x)\wedge\forall y(F(y)\rightarrow\neg A(x, y)))$ e) $\exists x(F(x)\wedge\forall y(S(y)\rightarrow\neg A(y, x)))$
f) $\forall y(F(y)\rightarrow\exists x(S(x)\vee A(x, y)))$ g) $\exists x(F(x)\wedge\forall y((F(y)\wedge(y\neq x))\rightarrow A(x, y)))$
h) $\exists x(S(x)\wedge\forall y(F(y)\rightarrow\neg A(y, x)))$

13. a) $\neg M(\text{Chou}, \text{Koko})$ b) $\neg M(\text{Arlene}, \text{Sarah})\wedge\neg T(\text{Arlene}, \text{Sarah})$
c) $\neg M(\text{Deborah}, \text{Jose})$ d) $\forall xM(x, \text{Ken})$ e) $\forall x\neg T(x, \text{Nina})$
f) $\forall x((T, x, \text{Avi})\vee M(x, \text{Avi}))$ g) $\exists x\forall y(y\neq x\rightarrow M(x, y))$
h) $\exists x\forall y(y\neq x\rightarrow(M(x, y)\vee T(x, y)))$ i) $\exists x\exists y(x\neq y\wedge M(x, y)\wedge M(y, x))$
j) $\exists xM(x, x)$ k) $\exists x\forall y(x\neq y\rightarrow(\neg M(x, y)\wedge\neg T(y, x)))$
l) $\forall x(\exists y(x\neq y\wedge(M(y, x)\vee T(y, x))))$ m) $\exists x\exists y(x\neq y\wedge M(x, y)\wedge T(y, x))$
n) $\exists x\exists y(x\neq y\wedge\forall z((z\neq x\wedge z\neq y)\rightarrow(M(x, z)\vee M(y, z)\vee T(x, z)\vee T(y, z))))$

15. a) $\forall xP(x)$，其中 $P(x)$为“x需要一门离散数学课”，论域是所有计算机科学的学生集合。
b) $\exists xP(x)$，其中 $P(x)$为“x有台 PC(计算机)”，论域是班上所有学生的集合。
c) $\forall x\exists yP(x, y)$，其中 $P(x, y)$为“x选修y”，x的论域是班上所有学生的集合，y的论域是计算机科学课程的集合。
d) $\exists x\exists yP(x, y)$，其中 $P(x, y)$以及 x、y的论域同 c)。
e) $\forall x\forall yP(x, y)$，其中 $P(x, y)$为“x去过y”，x的论域是班上所有学生的集合，y的论域是校园内所有楼房的集合。
f) $\exists x\exists y\forall z(P(z, y)\rightarrow Q(x, z))$，其中 $P(z, y)$为“z在y中”，$Q(x, z)$为“x去过z”，其中x的论域是班上所有学生的集合，y的论域是校园内所有楼房的集合，z的论域则是房间的集合。
g) $\forall x\forall y\exists z(P(z, y)\wedge Q(x, z))$，其中 P、Q及x、y、z的论域与 f)相同。

17. a) $\forall u\exists m(A(u, m)\wedge\forall n(n\neq m\rightarrow\neg A(u, n)))$，其中 $A(u, m)$意味着用户u已经访问过邮箱m。
b) $\exists p\forall e(H(e)\wedge S(p, \text{运行}))\rightarrow S(\text{内核}, \text{正确运行})$，其中 $H(e)$意味着错误条件e有效，$S(x, y)$意味着x的状态是y。
c) $\forall u\forall s(E(s, .\text{edu})\rightarrow A(u, s))$，其中 $E(s, x)$意味着站点s有扩展x，$A(u, s)$意味用户u可访问站点s。
d) $\exists x\exists y(x\neq y\wedge\forall z((\forall sM(z, s))\leftrightarrow(z=x\vee z=y)))$，其中 $M(a, b)$意味着系统a监控远程服务器b。

19. a) $\forall x\forall y((x<0)\wedge(y<0)\rightarrow(x+y<0))$ b) $\neg\forall x\forall y((x>0)\wedge(y>0)\rightarrow(x-y>0))$
c) $\forall x\forall y(x^2+y^2\geqslant(x+y)^2)$ d) $\forall x\forall y(|xy|=|x||y|)$

21. $\forall x\exists a\exists b\exists c\exists d((x>0)\rightarrow x=a^2+b^2+c^2+d^2)$，其中论域是全体整数。

23. a) $\forall x\forall y((x<0)\wedge(y<0)\rightarrow(xy>0))$ b) $\forall x(x-x=0)$
c) $\forall x\exists a\exists b(a\neq b\wedge\forall c(c^2=x\leftrightarrow(c=a\vee c=b)))$ d) $\forall x((x<0)\rightarrow\neg\exists y(x=y^2))$

25. a)对于实数，存在乘法的等式。
b)两个负实数的积总是正实数。
c)存在实数x和y使得$x^2>y$但$x<y$。
d)实数在加法操作下是有界的。

27. a) True b) True c) True d) True e) True

f)False g)False h)True i)False

29. a)$P(1,1)\wedge P(1,2)\wedge P(1,3)\wedge P(2,1)\wedge P(2,2)\wedge P(2,3)\wedge P(3,1)\wedge P(3,2)\wedge P(3,3)$
b)$P(1,1)\vee P(1,2)\vee P(1,3)\vee P(2,1)\vee P(2,2)\vee P(2,3)\vee P(3,1)\vee P(3,2)\vee P(3,3)$
c)$(P(1,1)\wedge P(1,2)\wedge P(1,3))\vee(P(2,1)\wedge P(2,2)\wedge P(2,3))\vee(P(3,1)\wedge P(3,2)\wedge P(3,3))$
d)$(P(1,1)\vee P(2,1)\vee P(3,1))\wedge(P(1,2)\vee P(2,2)\vee P(3,2))\wedge(P(1,3)\vee P(2,3)\vee P(3,3))$

31. a)$\exists x\forall y\exists z\neg T(x,y,z)$ b)$\exists x\forall y\neg P(x,y)\wedge\exists x\forall y\neg Q(x,y)$
c)$\exists x\forall y(\neg P(x,y)\vee\forall z\neg R(x,y,z))$ d)$\exists x\forall y(P(x,y)\wedge\neg Q(x,y))$

33. a)$\exists x\exists y\neg P(x,y)$ b)$\exists y\forall x\neg P(x,y)$
c)$\exists y\exists x(\neg P(x,y)\wedge\neg Q(x,y))$ d)$(\forall x\forall yP(x,y))\vee(\exists x\exists y\neg Q(x,y))$
e)$\exists x(\forall y\exists z\neg P(x,y,z)\vee\forall z\exists y\neg P(x,y,z))$

35. a)班上有这样的同学，在任何两门不同的数学课中，要么这两门都没选，要么他选了这两门。
b)每个人要么已访问过利比亚，要么从未访问过利比亚以外的任何国家。
c)有人已攀登过喜玛拉雅山的每座山峰。
d)有个人既没和 Kevin Bacon 出演同一部电影，也没和任何与 Kevin Bacon 出演过电影的人出演同一部电影。

37. a)$x=2,\ y=-2$ b)$x=-4$ c)$x=17,\ y=-1$

39. $\forall x\forall y\forall z((x\cdot y)\cdot z=x\cdot(y\cdot z))$

41. a)True b)False c)True

43. $\neg(\exists x\forall yP(x,y))\leftrightarrow\forall x(\neg\forall yP(x,y))\leftrightarrow\forall x\exists y\neg P(x,y)$

45. a)假定$\forall xP(x)\wedge\exists xQ(x)$为真，那么对所有 x，$P(x)$为真，而且有个元素 y 使$Q(y)$为真。因为$P(x)\wedge Q(y)$对所有 x 成真，而且有 y 使$Q(y)$成真，$\forall x\exists y(P(x)\wedge Q(y))$为真。反过来，假定第 2 个命题为真。令 x 为论域中的一个元素。有元素 y 使$Q(y)$为真，所以$\exists xQ(x)$为真。由于$\forall xP(x)$也为真，所以第 1 个命题也为真。
b)假定$\forall xP(x)\vee\exists xQ(x)$为真。那么要么对所有 x，$P(x)$为真；要么有 y，使$Q(y)$为真。在第 1 种情况下，$P(x)\vee Q(y)$对所有 x 成真，所以$\forall x\exists y(P(x)\vee Q(y))$为真。在第 2 种情况下，对特定的 y，$Q(y)$成真，所以 $P(x)\vee Q(y)$对所有 x 成真，从而$\forall x\exists y(P(x)\vee Q(y))$成真。反过来，假定第 2 个命题为真。如果 $P(x)$对所有 x 为真，那么第 1 个命题为真；否则 $P(x)$对某个 x 为假，对这一 x 必有 y 使 $P(x)\vee Q(y)$成真。因此 $Q(y)$必成真，于是$\exists yQ(y)$为真。这样第 1 个命题为真。

47. 我们将要说明，在表达式的子表达式均可转换为前束范式(PNF)的情况下，怎样把这个表达式转换为前束范式。这样从内向外就可以用这一方法把所有表达式转换为前束范式。(要使论证形式化，就得使用集合上的数学归纳法，这一方法在 3.4 节讨论。)由 1.2 节的练习 39 可知，可以假定命题中只使用逻辑联结符$\vee$和$\neg$。注意不带量词的命题已经是前束范式的形式。(这是论证的基本情况。)现在假定命题是 $QxP(x)$形的，其中 Q 为量词。因为 $P(x)$是比原命题短的表达式，它可以转换为 PNF。Qx 后面跟上这 PNF 仍是 PNF，它等价于原命题。其次，假定命题为 P。如果 P 已经是 PNF 形式的，我们可以用表 1-17 的等价关系把否定符号移到所有量词内层。最后，假定命题为 $P\vee Q$形的，其中 P 和 Q 均为 PNF。如果只有 P 和 Q 中的一个含量词，那么用 1.3 节练习 44 可以把量词移到它们两个的前面。如果 P 和 Q 均含量词，可以用 1.3 节练习 43、练习 44 或 45b)重写 $P\vee Q$，使两个量词都出现在形为 $R\vee S$ 的命题之前，然后再把$R\vee S$ 转换为 PNF。

49. $\exists xP(x)\wedge\forall x\forall y((P(x)\wedge P(y))\rightarrow x=y)$

51. $\forall L\exists\epsilon\forall N\exists n(n>N\wedge|a_n-L|\geqslant\epsilon)$

53. $\forall\epsilon(\forall N\exists n(n>N\wedge a_n>L-\epsilon)\wedge\exists N\forall n(n>N\rightarrow a_n\leqslant L+\epsilon))$

1.5 节

1. a)附加 b)化简 c)假言推理 d)拒取式 e)假言三段论

3. 设 w 是“兰迪努力工作”，设 d 是“兰迪是个笨小子”，设 j 是“兰迪将得到这份工作”。前提是 w，$w \to d$ 和 $d \to \neg j$。用假言推理和前两个前提得出 d。用假言推理和最后一个前提得出 $\neg j$，它就是所需要的结论：“兰迪将得不到这份工作”。
5. 用全称量词消去来得出“若苏格拉底是人，则苏格拉底是要死的”。用假言推理来得出苏格拉底是要死的。
7. a)有效结论是“我周二没有休假”，“我周四休过假”，“周四下过雨”。
   b)“我没有吃辣的食物而且没有打雷”是有效结论。
   c)“我是聪明的”是有效结论。
   d)“拉尔夫不是主修计算机科学的”是有效结论。
   e)“你购买许多东西对美国有利而且对你有利”是有效结论。
   f)“老鼠啃咬它们的食物”和“兔子不是鼠类”是有效结论。
9. a)设 $c(x)$ 是“x 是在这个班里”，$j(x)$ 是“x 知道如何用 JAVA 编写程序”，$h(x)$ 是“x 可以找到高薪工作”。前提是 c(道格)，j(道格)，$\forall x(j(x) \to h(x))$。用全称量词消去和最后一个前提得出 j(道格)$\to h$(道格)。对这个结论和第 2 个前提用假言推理得出 h(道格)。用合取和第 1 个前提得出 c(道格)$\wedge h$(道格)。最后，用存在量词引入得出所需要的结论 $\exists x(c(x) \wedge h(x))$。
   b)设 $c(x)$ 是“x 是在这个班里”，$w(x)$ 是“x 喜欢鲸鱼观察”，$p(x)$ 是“x 关心海洋污染”。前提是 $\exists x(c(x) \wedge w(x))$ 和 $\forall x(w(x) \to p(x))$。根据第 1 个前提，对一个具体的人 y 来说有 $c(y) \wedge w(y)$。化简得出 $w(y)$。用第 2 个前提和全称量词消去得出 $w(y) \to p(y)$。用假言推理得出 $p(y)$，再用合取得出 $c(y) \wedge p(y)$。最后，用存在量词引入得出所需要的结论 $\exists x(c(x) \wedge p(x))$。
   c)设 $c(x)$ 是“x 是在这个班里”，$p(x)$ 是“x 拥有一台个人电脑”，$w(x)$ 是“x 会使用字处理程序”。前提是 c(齐克)，$\forall x(c(x) \to p(x))$ 和 $\forall x(p(x) \to w(x))$。用第 2 个前提和全称量词消去得出 c(齐克)$\to p$(齐克)。用第 1 个前提和假言推理得出 p(齐克)。用第 3 个前提和全称量词消去得出 p(齐克)$\to w$(齐克)。最后，用假言推理得出所需要的结论 w(齐克)。
   d)设 $j(x)$ 是“x 是在新泽西”，$f(x)$ 是“x 住在距离大海 50 英里之内”，$s(x)$ 是“x 见过大海”。前提是 $\forall x(j(x) \to f(x))$ 和 $\exists x(j(x) \wedge \neg s(x))$。第 2 个前提和存在量词消去意味着对一个具体的人 y 来说有 $j(y) \wedge \neg s(y)$。化简得出对这个人 y 来说有 $j(y)$。用全称量词消去和第 1 个前提得出 $j(y) \to f(y)$，再用假言推理得出 $f(y)$。用化简从 $j(y) \wedge \neg s(y)$ 得出 $\neg s(y)$。所以用合取得出 $f(y) \wedge \neg s(y)$。最后，用存在量词引入得出所需要的结论 $\exists x(f(x) \wedge \neg s(x))$。
11. a)正确，使用全称量词消去和假言推理　　b)无效，肯定结论谬误
    c)无效，否定前提谬误　　d)正确，使用全称量词消去和假言推理
13. a)肯定结论谬误　　b)回避问题谬误
    c)使用假言推理的有效论证　　d)使用析取三段论的有效论证
    e)否定前提谬误
15. 我们知道存在一些 x 使得 $H(x)$ 为真，但不能得出结论 Lola 是这样的 x。
17. 这个命题平凡地为真，因为 0 不是正整数。空证明。
19. $P(1)$ 为真，因为 $(a+b)^1 = a+b \geq a^1 + b^1 = a+b$。直接证明。
21. a)假定 n 是奇数，所以对某个整数 k 来说有 $n=2k+1$。于是 $n^3+5=2(4k^3+6k^2+3k+3)$。因为 n^3+5 是 2 乘以某个整数，所以它是偶数。
    b)假设 n^3+5 是奇数而且 n 是奇数。因为 n 是奇数并且两个奇数之积是奇数，所以 n^2 是奇数而且 n^3 也是奇数。但是这样一来 $5=(n^3+5)-n^3$ 应当是偶数，因为它是两个奇数之差。因此，n^3+5 和 n 都是奇数这个假设是错误的。
23. 设 $n=2k+1$ 和 $m=2l+1$ 都是奇数，则 $n+m=2(k+l+1)$ 是偶数。
25. 假设 r 是有理数，i 是无理数，$s=r+i$ 是有理数。则根据练习 10，$s+(-r)=i$ 是有理数，矛盾。
27. 因为 $\sqrt{2} \cdot \sqrt{2} = 2$ 是有理数，且 $\sqrt{2}$ 是无理数，所以两个无理数之积不一定是无理数。

29. 间接证明：如果 $1/x$ 是有理数，则由定义 $1/x=p/q$，p 和 q 是整数且 $q\neq 0$。由于 $1/x$ 不能为 0(如果为零，则将两边同乘 x 会得出 $1=x\cdot 0$ 的矛盾)，可知 $p\neq 0$。由代数和运算规则，$x=1/(1/x)=1/(p/q)=q/p$。因此 x 可以写成两个整数的商，分母非零。因此由定义，x 是有理数。

31. 如果在一个星期的每一天中有 9 天(或更少)，那么最多有 $9\times 7=63$ 天。但我们选择了 64 天。这一矛盾证明至少有 10 天指向一个星期的同一天。

33. 若 $x\leqslant y$，则 $\max(x, y)+\min(x, y)=y+x=x+y$。若 $x\geqslant y$，则 $\max(x, y)+\min(x, y)=x+y$。因为这是仅有的两种情况，所以不等式总是成立。

35. 存在四种情形。情形 1：$x\geqslant 0$ 而且 $y\geqslant 0$，于是 $|x|+|y|=x+y=|x+y|$。情形 2：$x<0$ 而且 $y<0$，于是 $|x|+|y|=-x+(-y)=-(x+y)=|x+y|$，因为 $x+y<0$。情形 3：$x\geqslant 0$ 而且 $y<0$，于是 $|x|+|y|=x+(-y)$。若 $x\geqslant -y$，则 $|x+y|=x+y$。但是因为 $y<0$，$-y>y$，所以 $|x|+|y|=x+(-y)>x+y=|x+y|$。若 $x<-y$，则 $|x+y|=-(x+y)=-x+(-y)$。但是因为 $x\geqslant 0$，$-x\leqslant x$，所以 $|x|+|y|=x+(-y)\geqslant -x+(-y)=|x+y|$。情形 4：$x<0$ 而且 $y\geqslant 0$。交换 x 与 y，就等同于情形 3。

37. 不失一般性，我们可以假设 n 是非负的，因为整数的四次方及其负数的四次方是相同的。根据练习 36 的提示，我们将任意正整数 n 除以 10，得到商 k 和余数 l，因此 $n=10k+l$。l 是 0 和 9 之间的整数。然后我们计算这 10 种情况下的 n^4。我们得到如下值，其中 X 是 10 的倍数的某个整数，具体值不必关心。

$$
\begin{aligned}
(10k+0)^4 &= 10\,000k^4 = 10\,000k^4+0,\\
(10k+1)^4 &= 10\,000k^4+X\cdot k^3+X\cdot k^2+X\cdot k+1,\\
(10k+2)^4 &= 10\,000k^4+X\cdot k^3+X\cdot k^2+X\cdot k+16,\\
(10k+3)^4 &= 10\,000k^4+X\cdot k^3+X\cdot k^2+X\cdot k+81,\\
(10k+4)^4 &= 10\,000k^4+X\cdot k^3+X\cdot k^2+X\cdot k+256,\\
(10k+5)^4 &= 10\,000k^4+X\cdot k^3+X\cdot k^2+X\cdot k+625,\\
(10k+6)^4 &= 10\,000k^4+X\cdot k^3+X\cdot k^2+X\cdot k+1296,\\
(10k+7)^4 &= 10\,000k^4+X\cdot k^3+X\cdot k^2+X\cdot k+2401,\\
(10k+8)^4 &= 10\,000k^4+X\cdot k^3+X\cdot k^2+X\cdot k+4096,\\
(10k+9)^4 &= 10\,000k^4+X\cdot k^3+X\cdot k^2+X\cdot k+6561。
\end{aligned}
$$

由于 X 是 10 的倍数，因此对应项不会对结果的个位数产生影响。因此，个位数分别是 0，1，6，1，6，5，6，1，6，1，所以个位数总是 0，1，5 或 6。

39. 首先，假定 n 是奇数，所以对某个整数 k 来说 $n=2k+1$。于是 $5n+6=5(2k+1)+6=10k+11=2(5k+5)+1$。因此 $5n+6$ 是奇数。为了证明逆命题，假定 n 是偶数，所以对某个整数 k 来说有 $n=2k$。于是 $5n+6=10k+6=2(5k+3)$，所以 $5n+6$ 是偶数。因此 n 是奇数当且仅当 $5n+6$ 是奇数。

41. 这个命题为真。假定 m 既不是 1 也不是 -1。则 mn 有大于 1 的因子 m。在另一方面，$mn=1$，且 1 没有这样的因子。因此 $m=1$ 或 $m=-1$。在第 1 种情形里 $n=1$，而在第 2 种情形里 $n=-1$，因为 $n=1/m$。

43. 我们首先证明 x 为偶数时，三个命题等价。若 x 是偶数，则对某个整数 k 有 $x=2k$。因此 $3x+2=3\cdot 2k+2=6k+2=2(3k+1)$。这是偶数，因为它形如 $2t$，其中 $t=3k+1$。类似地，$x+5=2k+5=2k+4+1=2(k+2)+1$，因此 $x+5$ 是奇数；$x^2=(2k)^2=2(2k^2)$，因此 x^2 是偶数。现在用间接证明考虑另一种情况，假设 x 是奇数，则可将 x 写为 $x=2k+1$。此时 $3x+2=3(2k+1)+2=6k+5=2(3k+2)+1$，它是奇数，因为它可以写成 $2t+1$ 的形式，其中 $t=3k+2$。类似地，$x+5=2k+1+5=2(k+3)$，因此 $x+5$ 是偶数。x^2 是奇数已在例 14 中证明过了。

45. 我们给出(i)→(ii)，(ii)→(i)，(i)→(iv)及(iii)→(i)的间接证明。首先，假设 $3x+2$ 是有理数，而 $3x+2=p/q$，p 和 q 是整数且 $q\neq 0$。则可以写 $x=((p/q)-2)/3=(p-2q)/(3q)$，其中 $3q\neq 0$。这证明 x 是有理数。对第二个蕴含，假设 x 是有理数，即 $x=p/q$，p 和 q 是整数且 $q\neq 0$。则 $3x+2=(3p+2q)/q$，

$q\neq0$。这证明 $3x+2$ 是有理数。对第三个蕴含，假设 $x/2$ 是有理数，即 $x/2=p/q$，p 和 q 是整数且 $q\neq0$。则可以写 $x=2p/q$，其中 $q\neq0$。这证明 x 是有理数。对第四个蕴含，假设 x 是有理数。即 $x=p/q$，p 和 q 是整数且 $q\neq0$，则可以写 $x/2=p/2q$；其中 $2q\neq0$。这证明 $x/2$ 是有理数。

47. 不正确

49. 10 001，10 002，…，10 100 是不完全平方的，因为 $100^2=10\,000$ 且 $101^2=10\,201$。构造性的。

51. $8=2^3$ 而 $9=3^2$

53. a)这个命题断言存在具有特定性质的 x。如果令 $y=x$，则可见 $P(x)$ 为真。如果 y 不是 x，则 $P(x)$ 不为真，因此 x 是使 P 为真的唯一元素。

b)第一个子句表明有某个元素使 P 为真；第二个子句表明每当有两个元素使 P 为真，这两个元素是相同的。合在一起就表示 P 仅被一个元素所满足。

c)这个命题断言存在 x 使 P 为真，且只要我们发现有元素使 P 为真，这个元素就是 x。换句话说，x 是使 P 为真的唯一元素。

55. 等式 $|a-c|=|b-c|$ 等价于 $a-c=b-c$ 或 $a-c=-b+c$ 的析取。第一个式子等价于 $a=b$，这与假设矛盾，因此原式等价于 $a-c=-b+c$。在等式两边加上 $b+c$ 并除以 2 可得 $c=(a+b)/2$。因此有唯一解。而且，因为奇整数 a 和 b 的和是偶数，所以 $c=(a+b)/2$ 是整数。

57. 我们要求解 $n=(k-2)+(k+3)$ 中的 k。使用一般的代数规则，可知此式等价于 $k=(n-1)/2$。换句话说，这是唯一可使等式为真的 k 值。因为 n 为奇数，$n-1$ 为偶数，因此 k 是整数。

59. 若 x 是整数，则可取 $n=x$ 且 $\epsilon=0$。在这种情况下无其他解，因为如果整数 n 大于 x，则 n 至少是 $x+1$，这将使 $\epsilon\geqslant1$。如果 x 不是整数，则向上取整，称此整数为 n。令 $\epsilon=n-x$。显然 $0\leqslant\epsilon<1$，对这个 n，这是唯一合适的 ϵ，且 n 不能再大，因为 ϵ 限制为小于 1。

61. 令 p 为“下雨了”；令 q 为“Yvette 带伞了”；令 r 为“Yvette 被淋湿了”。三个前提分别是 $\neg p\vee q$，$\neg q\vee\neg r$ 和 $p\vee\neg r$。对前两者进行消解得出 $\neg p\vee\neg r$。将此式与第三个式子消解得出 $\neg r$，题目得证。

63. 假设这个命题是可满足的。对前两个子句使用消解可得出 $q\vee q$，换句话说，q 必须为真。对后两个子句运用消解可得 $\neg q\vee\neg q$，换句话说，$\neg q$ 必须为真。这是矛盾，因此命题不是可满足的。

65. 令 $x=2$ 且 $y=\sqrt{2}$。若 $x^y=2^{\sqrt{2}}$ 是无理数，即完成证明。若不是，则令 $x=2^{\sqrt{2}}$ 且 $y=\sqrt{2}/4$。则 $x^y=(2^{\sqrt{2}})^{\sqrt{2}/4}=2^{\sqrt{2}\cdot(\sqrt{2})/4}=2^{1/2}=\sqrt{2}$。

67. 假设 $p_1\rightarrow p_4\rightarrow p_2\rightarrow p_5\rightarrow p_3\rightarrow p_1$。为了证明这些命题中的一个蕴含着其余的任意一个，只需要重复使用假言三段论。

69. 每个放在棋盘上的多米诺骨牌都恰好覆盖一个白格和一个黑格，因此一组多米诺骨牌恰好覆盖同样数目的白格和黑格。因为删除对角上的格子产生出黑格比白格多两个或白格比黑格多两个的棋盘，所以没有一组多米诺骨牌可以覆盖删除了对角上的格子的棋盘。

71. 将给出归谬证明。假定 $a_1, a_2, \cdots, a_n$ 都小于 A，其中 A 是这些数的平均值。则 $a_1+a_2+\cdots+a_n<nA$。两边都除以 n 就证明 $A=(a_1+a_2+\cdots+a_n)/n<A$，矛盾。

73. 将通过证明(i)蕴含(ii)、(ii)蕴含(iii)、(iii)蕴含(iv)的(iv)蕴含(i)来证明这四个命题都是等价的。首先，假设 n 是偶数。则对某个整数 k 来说 $n=2k$。于是 $n+1=2k+1$，所以 $n+1$ 是奇数。则对某个整数 k 来说 $3n+1=2k+1$。于是 $3n=(2k+1)-1=2k$，所以 $3n$ 是偶数。这证明(i)蕴含(ii)。其次，假设 $n+1$ 是奇数。则对某个整数 k 来说 $n+1=2k+1$。于是 $3n+1=2n+(n+1)=2(n+k)+1$，这说明 $3n+1$ 是奇数，证明(ii)蕴含(iii)。再次，假设 $3n+1$ 是奇数。这证明(iii)蕴含(iv)。最后，假设 n 不是偶数。则 n 是奇数，所以对某个整数 k 来说 $n=2k+1$。于是 $3n=3(2k+1)=6k+3=2(3k+1)+1$，所以 $3n$ 是奇数。这样完成了(iv)蕴含(i)的间接证明。

75. 根据第二个前提，存在某个不喝咖啡的狮子。设里欧是这样的动物。通过化简知道里欧是狮子。通过假言推理，从第 1 个前提知道里欧是凶猛。因此里欧是凶猛的并且是不喝咖啡的。根据存在量词的定义，存在着不喝咖啡的凶猛的动物，即某些不喝咖啡的凶猛动物。

77. 有效。

1.6 节

1. a) $\{-1, 1\}$　　b) $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$
   c) $\{0, 1, 4, 9, 16, 25, 36, 49, 64, 81\}$　　d) $\varnothing$
3. a)相等　　b)不等　　c)不等
5. a)是　　b)不是　　c)是
   d)不是　　e)不是　　f)不是
7. a)False　　b)False　　c)False
   d)True　　e)False　　f)False　　g)True
9. a)True　　b)True　　c)False
   d)True　　e)True　　f)False
11. 假定 $x\in A$。由于 $A\subseteq B$，则 $x\in B$。由于 $B\subseteq C$，则 $x\in C$。由 $x\in A$ 导出 $x\in C$，所以 $A\subseteq C$。
13. a)1　　b)1　　c)2　　d)3
15. a) $\{\varnothing, \{a\}\}$　　b) $\{\varnothing, \{a\}, \{b\}, \{a, b\}\}$
    c) $\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}$
17. a)8　　b)16　　c)2
19. a) $\{(a, y), (b, y), (c, y), (d, y), (a, z), (b, z), (c, z), (d, z)\}$
    b) $\{(y, a), (y, b), (y, c), (y, d), (z, a), (z, b), (z, c), (z, d)\}$
21. 三元组 (a, b, c) 的集合，其中 a 是航线，b 和 c 是城市。
23. $\varnothing\times A=\{(x, y)\mid x\in\varnothing\wedge y\in A\}=\varnothing=\{(x, y)\mid x\in A\wedge y\in\varnothing\}=A\times\varnothing$
25. mn
27. a)实数的平方绝不会是 -1。为真。　　b)存在某个整数它的平方等于 2。为假。
    c)每个整数的平方都是正的。为假。　　d)存在某个实数等于它自身的平方。为真。
29. 我们必须证明 $\{\{a\}, \{a, b\}\}=\{\{c\}, \{c, d\}\}$ 成立的充分必要条件是 $a=c$ 且 $b=d$。显然条件是充分的。为证必要性，假定这两个集合相等。首先考虑 $a\neq b$ 的情况，这时 $\{\{a\}, \{a, b\}\}$ 恰含两个元素，其中的一个包含一个元素。这样 $\{\{c\}, \{c, d\}\}$ 也必须有同样的性质，所以 $c\neq d$，且 $\{c\}$ 是两个元素中只含一个元素的那一个。于是 $\{a\}=\{c\}$，也就是 $a=c$。含两个元素的集合 $\{a, b\}$ 和 $\{c, d\}$ 也必须相等。由于 $a=c$，且 $a\neq b$，所以 $b=d$。其次假定 $a=b$。于是 $\{\{a\}, \{a, b\}\}=\{\{a\}\}$，这是只含一个元素的集合。于是 $\{\{c\}, \{c, d\}\}$ 也只含一个元素，这只有在 $c=d$ 时才有可能，此时该集合为 $\{\{c\}\}$。从而 $a=c$ 且 $b=d$。
31. 令 $S=\{a_1, a_2, \cdots, a_n\}$。把 S 的每个子集 S' 都用长度为 n 的位串表示，位串中第 i 位为 1 的充分必要条件是 $a_i\in S$。为产生 S 的所有子集，列出长度为 n 的所有 2^n 个位串(例如，按增序)，再写下相应的子集。

1.7 节

1. a)住处离校不超过 1 英里且走路上学的学生集合。
   b)住处离校不超过 1 英里或走路上学的学生之集合。
   c)住处离校不超过 1 英里但不走路上学的学生集合。
   d)走路上学但住处离校超过 1 英里的学生集合。
3. a) $\{0, 1, 2, 3, 4, 5, 6\}$　　b) $\{3\}$　　c) $\{1, 2, 4, 5\}$　　d) $\{0, 6\}$
5. $\overline{\overline{A}}=\{x\mid \neg(x\in\overline{A})\}=\{x\mid \neg(\neg x\in A)\}=\{x\mid x\in A\}=A$。
7. a) $A\cup B=\{x\mid x\in A\vee x\in B\}$
   $=\{x\mid x\in B\vee x\in A\}=B\cup A$
   b) $A\cap B=\{x\mid x\in A\wedge x\in B\}$
   $=\{x\mid x\in B\wedge x\in A\}=B\cap A$

9. 假定 $x\in A\cap(A\cup B)$。则由交集的定义 $x\in A$ 且 $x\in A\cup B$。由于 $x\in A$，我们就已证明了等式左边是右边的子集。反之，令 $x\in A$。则由并集的定义，$x\in A\cup B$。因此由交集的定义 $x\in A\cap(A\cup B)$，所以右边是左边的子集。

11. a) $x\in\overline{(A\cup B)}\equiv x\notin(A\cup B)\equiv\neg(x\in A\vee x\in B)$

$\equiv\neg(x\in A)\wedge\neg(x\in B)\equiv x\notin A\wedge x\notin B\equiv x\in\bar{A}\wedge x\in\bar{B}\equiv x\in\bar{A}\cap\bar{B}$

b)

A	B	$A\cup B$	$\overline{(A\cup B)}$	$\bar{A}$	$\bar{B}$	$\bar{A}\cap\bar{B}$
1	1	1	0	0	0	0
1	0	1	0	0	1	0
0	1	1	0	1	0	0
0	0	0	1	1	1	1

13. a) $x\in\overline{A\cap B\cap C}\equiv x\notin A\cap B\cap C\equiv x\notin A\vee x\notin B\vee x\notin C\equiv x\in\bar{A}\vee x\in\bar{B}\vee x\in\bar{C}\equiv x\in\bar{A}\cup\bar{B}\cup\bar{C}$

b)

$A\ B\ C$	$A\cap B\cap C$	$\overline{(A\cap B\cap C)}$	$\bar{A}\ \bar{B}\ \bar{C}$	$\bar{A}\cup\bar{B}\cup\bar{C}$
1 1 1	1	0	0 0 0	0
1 1 0	0	1	0 0 1	1
1 0 1	0	1	0 1 0	1
1 0 0	0	1	0 1 1	1
0 1 1	0	1	1 0 0	1
0 1 0	0	1	1 0 1	1
0 0 1	0	1	1 1 0	1
0 0 0	0	1	1 1 1	1

15. 两边都等于 $\{x\mid x\in A\wedge x\notin B\}$。

17. a) $x\in A\cup(B\cup C)\equiv(x\in A)\vee(x\in(B\cup C))\equiv(x\in A)\vee(x\in B\vee x\in C)\equiv(x\in A\vee x\in B)\vee(x\in C)\equiv x\in(A\cup B)\cup C$

b) 只需将 a) 中 $\cup$ 替换成 $\cap$，$\vee$ 替换成 $\wedge$。

c) $x\in A\cup(B\cap C)\equiv(x\in A)\vee(x\in(B\cap C))\equiv(x\in A)\vee(x\in B\wedge x\in C)\equiv(x\in A\vee x\in B)\wedge(x\in A\vee x\in C)\equiv x\in(A\cup B)\cap(A\cup C)$

19. a) $\{4, 6\}$　b) $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$

c) $\{4, 5, 6, 8, 10\}$　d) $\{0, 2, 4, 5, 6, 7, 8, 9, 10\}$

21. a) $B\subseteq A$　b) $A\subseteq B$　c) $A\cap B=\varnothing$

d) 没什么可说，因为它总是成真。　e) $A=B$

23. $A\subseteq B\equiv\forall x(x\in A\to x\in B)\equiv\forall x(x\notin B\to x\notin A)\equiv\forall x(x\in\bar{B}\to x\in\bar{A})\equiv\bar{B}\subseteq\bar{A}$

25. 主修计算机科学而不主修数学，或主修数学而不主修计算机科学的所有学生的集合。

27. $(A\cup B)-(A\cap B)$ 的元素是属于 A 和 B 的并集却不属于 A 和 B 的交集的元素，也就是说它属于 A 或属于 B，但不同时属于 A 和 B。这正是元素属于 $A\oplus B$ 的含义。

29. a) $A\oplus A=(A-A)\cup(A-A)=\varnothing\cup\varnothing=\varnothing$　b) $A\oplus\varnothing=(A-\varnothing)\cup(\varnothing-A)=A\cup\varnothing=A$

c) $A\oplus U=(A-U)\cup(U-A)=\varnothing\cup\bar{A}=\bar{A}$　d) $A\oplus\bar{A}=(A-\bar{A})\cup(\bar{A}-A)=A\cup\bar{A}=U$

31. $B=\varnothing$

33. 是的。假定 $x\in A$ 但 $x\notin B$。如果 $x\in C$，那么 $x\notin A\oplus C$ 但 $x\in B\oplus C$，矛盾。如果 $x\notin C$，那么 $x\in A\oplus C$ 但 $x\notin B\oplus C$，矛盾。因此 $A\subseteq B$。类似地可知 $B\subseteq A$，从而 $A=B$。

35. 是的。

37. a) $\{1, 2, 3, \cdots, n\}$　b) $\{1\}$

39. a) A_n　b) $\{0, 1\}$

41. a) $\{1, 2, 3, 4, 7, 8, 9, 10\}$　b) $\{2, 4, 5, 6, 7\}$　c) $\{1, 10\}$

43. 如果第一个位串的第 i 位是 1 而第二个位串的第 i 位为 0，则两个集合之差的位串的第 i 位是 1；否则为 0。

45. a)11 1110 0000 0000 0000 0000 0000 ∨ 01 1100 1000 0000 0100 0101 0000 = 11 1110 1000 0000 0100 0101 0000，代表$\{a, b, c, d, e, g, p, t, v\}$

b)11 1110 0000 0000 0000 0000 0000 ∧ 01 1100 1000 0000 0100 0101 0000 = 01 1100 0000 0000 0000 0000 0000，代表$\{b, c, d\}$

c)(11 1110 0000 0000 0000 0000 0000 ∨ 00 0110 0110 0001 1000 0110 0110) ∧ (01 1100 1000 0000 0100 0101 0000 ∨ 00 1010 0010 0000 1000 0010 0111)
= 11 1110 0110 0001 1000 0110 0110 ∧ 01 1110 1010 0000 1100 0111 0111
= 01 1110 0010 0000 1000 0110 0110，代表$\{b, c, d, e, i, o, t, u, x, y\}$

d)11 1110 0000 0000 0000 0000 0000 ∨ 01 1100 1000 0000 0100 0101 0000 ∨ 00 1010 0010 0000 1000 0010 0111 ∨ 00 0110 0110 0001 1000 0110 0110 = 11 1110 1110 0001 1100 0111 0111，代表$\{a, b, c, d, e, g, h, i, n, o, p, t, u, v, x, y, z\}$

47. a)$\{1, 2, 3, \{1, 2, 3\}\}$　b)$\{\varnothing\}$
c)$\{\varnothing, \{\varnothing\}\}$　d)$\{\varnothing, \{\varnothing\}, \{\varnothing, \{\varnothing\}\}\}$

49. a)$\{3\cdot a, 3\cdot b, 1\cdot c, 4\cdot d\}$　b)$\{2\cdot a, 2\cdot b\}$
c)$\{1\cdot a, 1\cdot c\}$　d)$\{1\cdot b, 4\cdot d\}$　e)$\{5\cdot a, 5\cdot b, 1\cdot c, 4\cdot d\}$

51. $\overline{F}$ = {0.4 Alice, 0.1 Brian, 0.6 Fred, 0.9 Oscar, 0.5 Rita}
$\overline{R}$ = {0.6 Alice, 0.2 Brian, 0.8 Fred, 0.1 Oscar, 0.3 Rita}

53. $F \cap R$ = {0.4 Alice, 0.8 Brian, 0.2 Fred, 0.1 Oscar, 0.5 Rita}

1.8 节

1. a) $f(0)$无定义。　b)对 $x<0$，$f(x)$无定义。
c)对 x 的每个值都指派了两个值，$f(x)$不符合函数定义。

3. a)不是函数。　b)是函数。　c)不是函数。

5. a)整数集合。　b)非负偶数集合。
c)不超过 7 的非负整数集合。　d)整数平方的集合 = $\{0, 1, 4, 9, 16, \cdots\}$。

7. a) $\mathbf{Z}^+ \times \mathbf{Z}^+$；$\mathbf{Z}^+$　b)$\mathbf{Z}^+$；$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$
c)位串集合；$\mathbf{N}$　d)位串集合；$\mathbf{N}$

9. a)1　b)0　c)0　d)−1
e)3　f)−1　g)2　h)1

11. 只有 a)中的函数是。

13. 只有 a)和 d)中的函数是。

15. a)映上　b)不映上　c)映上
d)不映上　e)映上

17. a)$x\geqslant 0$ 时函数 $f(x)=3x+1$，$x<0$ 时函数 $f(x)=-3x+2$
b)$f(x)=|x|+1$
c)$x\geqslant 0$ 时 $f(x)=2x+1$，$x<0$ 时 $f(x)=-2x$
d)$f(x)=x^2+1$

19. a)是　b)不是　c)是　d)不是

21. 假定 f 严格递减。这意味着当 $x<y$ 时，$f(x)>f(y)$。要证明 g 严格递增，假定 $x<y$。则$g(x)=1/f(x)<1/f(y)=g(y)$。反过来，假定 g 严格递增。这意味着当 $x<y$ 时，$g(x)<g(y)$。要证明 f 严格递减，假定 $x<y$。则 $f(x)=1/g(x)>1/g(y)=f(y)$。

23. a) $f(S)=\{0, 1, 3\}$　b) $f(S)=\{0, 1, 3, 5, 8\}$
c) $f(S)=\{0, 8, 16, 40\}$　d) $f(S)=\{1, 12, 33, 65\}$

25. a)令 x 和 y 为 A 中不同的元素。由于 g 是一对一的，$g(x)$ 和 $g(y)$ 是 B 中不同的元素。又因为 f 是一对一的，$f(g(x))=(f\circ g)(x)$ 和 $f(g(y))=(f\circ g)(y)$ 是 C 中不同元素。因而 $f\circ g$ 是一对一的。
b)令 $y\in C$。因为 f 是映上的，对某个 $b\in B$，$y=f(b)$。又因为 g 是映上的，对某个 $x\in A$，$b=g(x)$。因此 $y=f(b)=f(g(x))=(f\circ g)(x)$。由此知 $f\circ g$ 是映上的。

27. 不是。例如，假定 $A=\{a\}$，$B=\{b, c\}$，$C=\{d\}$。令 $g(a)=b$，$f(b)=d$，$f(c)=d$，于是 f 和 $f\circ g$ 为映上的，但 g 不是。

29. $(f+g)(x)=x^2+x+3$，$(fg)(x)=x^3+2x^2+x+2$

31. f 是一对一的，因为 $f(x_1)=f(x_2)\equiv ax_1+b=ax_2+b\equiv ax_1=ax_2\equiv x_1=x_2$。$f$ 是映上的，因为 $f((y-b)/a)=y$，$f^{-1}(y)=(y-b)/a$。

33. 令 $f(1)=a$，$f(2)=a$，令 $S=\{1\}$，$T=\{2\}$。那么 $f(S\cap T)=f(\varnothing)=\varnothing$，但 $f(S)\cap f(T)=\{a\}\cap\{a\}=\{a\}$。

35. a) $\{x\mid 0\leqslant x<1\}$　b) $\{x\mid -1\leqslant x<2\}$　c) $\varnothing$

37. $f^{-1}(\overline{S})=\{x\in A\mid f(x)\notin S\}=\overline{\{x\in A\mid f(x)\in S\}}$
$=\overline{f^{-1}(S)}$。

39. 令 $x=\lfloor x\rfloor+\epsilon$，其中 ϵ 为实数，$0\leqslant\epsilon<1$。若 $\epsilon<\frac{1}{2}$，那么 $\lfloor x\rfloor-1<x-\frac{1}{2}<\lfloor x\rfloor$，所以 $\lceil x-1/2\rceil=\lfloor x\rfloor$ 且这是最接近 x 的整数。若 $\epsilon>\frac{1}{2}$，那么 $\lfloor x\rfloor<x-\frac{1}{2}<\lfloor x\rfloor+1$，所以 $\lceil x-1/2\rceil=\lfloor x\rfloor+1$ 且这是最接近 x 的整数。若 $\epsilon=\frac{1}{2}$，那么 $\lceil x-1/2\rceil=\lfloor x\rfloor$，这是围绕 x 且到 x 的距离相等的两个整数中较小的一个。

41. 把实数 x 写成 $\lfloor x\rfloor+\epsilon$，其中 ϵ 是实数，$0\leqslant\epsilon<1$。因为 $\epsilon=x-\lfloor x\rfloor$，所以 $0\leqslant-\lfloor x\rfloor<1$，由此直接得前两个不等式 $x-1<\lfloor x\rfloor$ 和 $\lfloor x\rfloor\leqslant x$。对另两个不等式，把 x 写成 $x=\lceil x\rceil-\epsilon'$，其中 $0\leqslant\epsilon'<1$。于是 $0\leqslant\lceil x\rceil-x<1$，从而立即得到想要的不等式。

43. a)若 $x<n$，由于 $\lfloor x\rfloor\leqslant x$，所以 $\lfloor x\rfloor<n$。假定 $x\geqslant n$，由底函数的定义知 $\lfloor x\rfloor\geqslant n$。这表明如果 $\lfloor x\rfloor<n$，那么 $x<n$。
b)若 $n<x$，那么由于 $x\leqslant\lceil x\rceil$，$n\leqslant\lceil x\rceil$。假定 $n\geqslant x$，由顶函数的定义知 $\lceil x\rceil\leqslant n$。这表明如果 $n<\lceil x\rceil$，那么 $n<x$。

45. 如果 n 为偶数，则对某个整数 k 有 $n=2k$。因此 $\lfloor n/2\rfloor=\lfloor k\rfloor=k=n/2$。
如果 n 是奇数，则对某个整数 k 有 $n=2k+1$。因此 $\lfloor n/2\rfloor=\left\lfloor k+\frac{1}{2}\right\rfloor=k=(n-1)/2$。

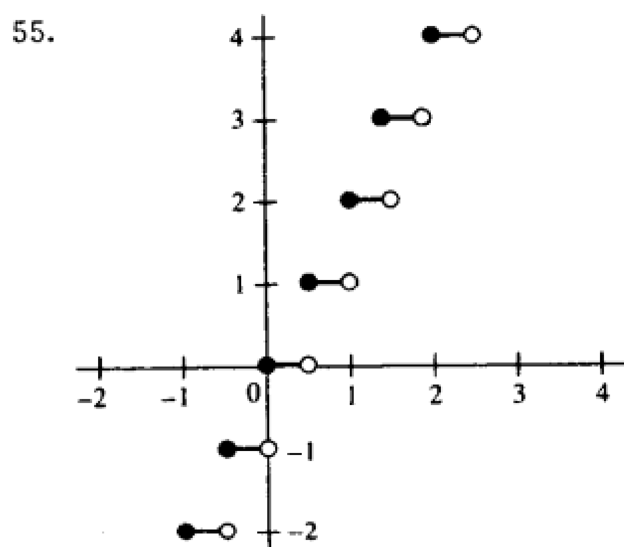
47. 假定 $x\geqslant 0$。等式左边是 $\lceil -x\rceil$ 而右边是 $\lfloor -x\rfloor$。如果 x 是整数，那么两边都等于 $-x$。否则令 $x=n+\epsilon$，其中 n 为自然数，ϵ 为实数，$0\leqslant\epsilon<1$。那么 $\lceil -x\rceil=\lceil -n-\epsilon\rceil=-n$，且 $\lfloor -x\rfloor=\lfloor -n+\epsilon\rfloor=-n$，等式成立。若 $x<0$，那么只要用 $-x$ 代替 x，即可得到想证明的等式。

49. $\lceil b\rceil-\lfloor a\rfloor-1$

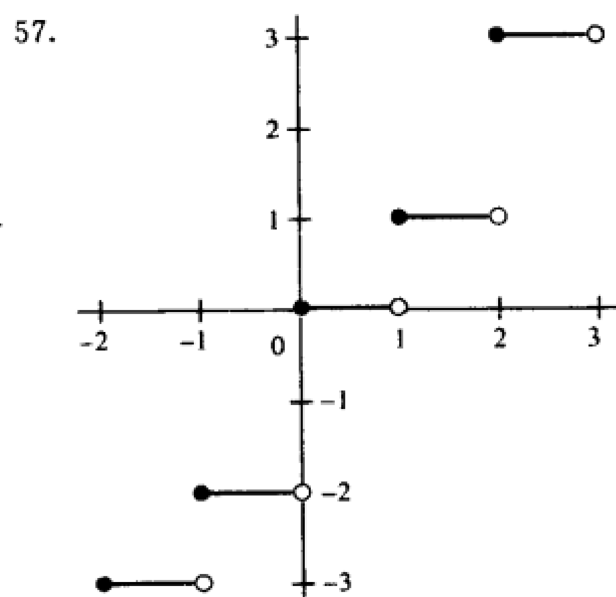
51. a) 1　b) 3　c) 126　d) 3600

53. a) 100　b) 256　c) 1030　d) 30 200

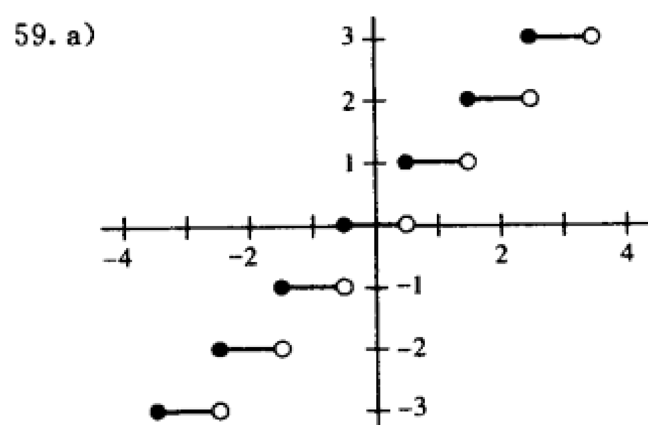
55.



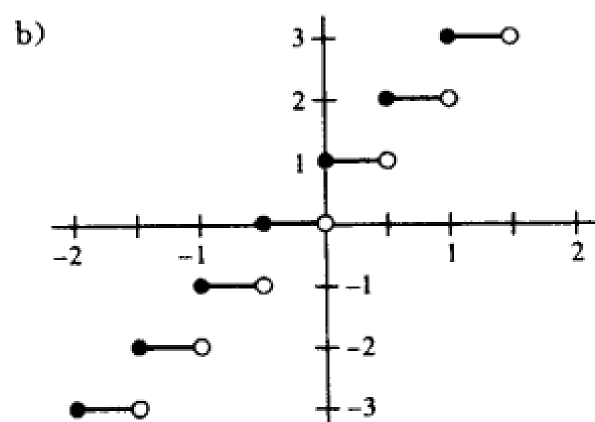
57.



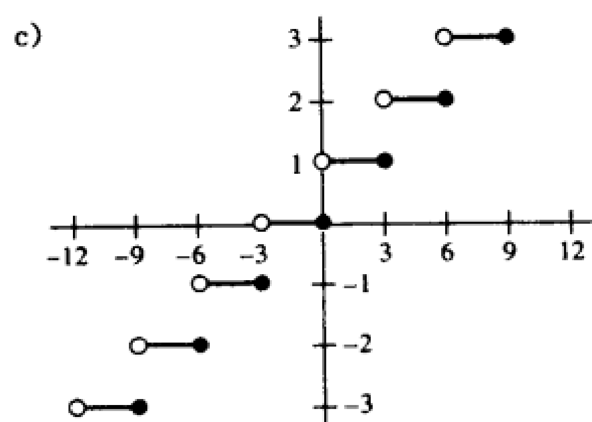
59. a)



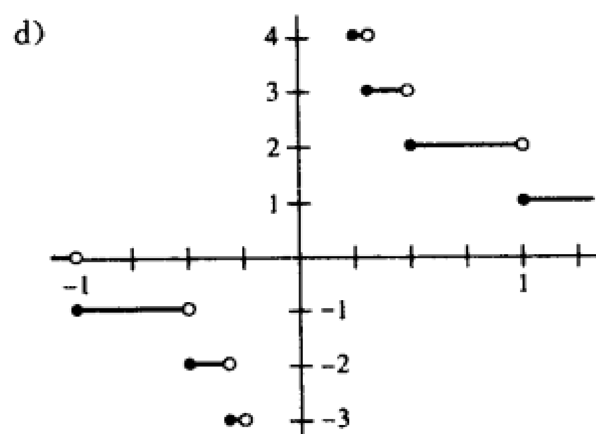
b)



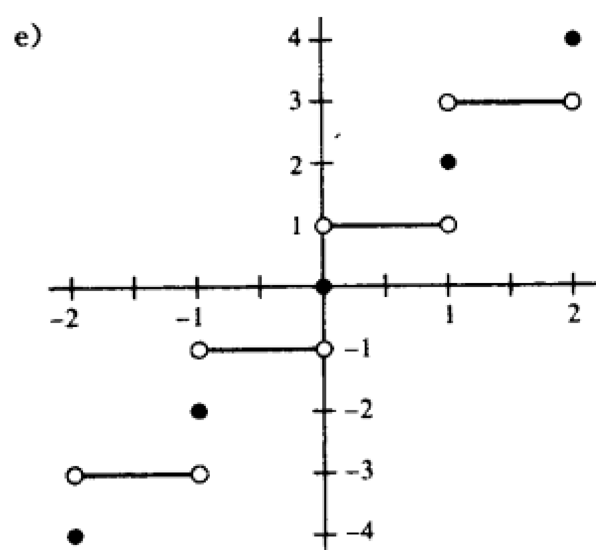
c)



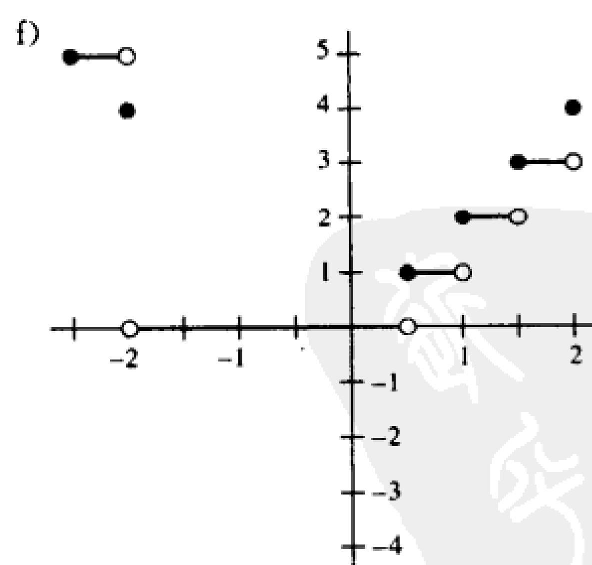
d)



e)



f)



g)见部分(a)

61. $f^{-1}(y)=(y-1)^{1/3}$

63. a) $f_{A\cap B}(x)=1 \equiv x\in A\cap B \equiv x\in A \wedge x\in B$

$\equiv f_A(x)=1 \wedge f_B(x)=1 \equiv f_A(x)f_B(x)=1$

b) $f_{A\cup B}(x)=1\equiv x\in A\cup B\equiv x\in A\vee x\in B$
$\equiv f_A(x)=1\vee f_B(x)=1\equiv f_A(x)+f_B(x)-f_A(x)f_B(x)=1$
c) $f_{\overline{A}}(x)=1\equiv x\in\overline{A}\Leftrightarrow x\notin A\equiv f_A(x)=0$
$\equiv 1-f_A(x)=1$
d) $f_{A\oplus B}(x)=1\equiv x\in A\oplus B\equiv(x\in A\wedge x\notin B)\vee$
$(x\notin A\wedge x\in B)\equiv f_A(x)+f_B(x)-2f_A(x)f_B(x)=1$

65. a)为真；因为$\lfloor x\rfloor$已经是整数，$\lceil\lfloor x\rfloor\rceil=\lfloor x\rfloor$。

b)为假；$x=\frac{1}{2}$是一个反例。

c)为真；如果 x 或 y 是整数，则使用1.8节表1-23的性质4b，差为0。如果 x 和 y 都不是整数，则$x=n+\epsilon$且$y=m+\delta$，其中 n 和 m 是整数，ϵ 和 δ 是小于1的正实数。则 $m+n<x+y<m+n+2$，因此$\lceil x+y\rceil$或者为$m+n+1$，或者为$m+n+2$。因此，给定的表达式或者为$(n+1)+(m+1)-(m+n+1)=1$，或者为$(n+1)+(m+1)-(m+n+2)=0$。

d)为假；$x=\frac{1}{4}$且 $y=3$ 是一个反例。

e)为假；$x=\frac{1}{2}$是一个反例。

67. a)如果 x 是正整数，则两边相等。因此假定 $x=n^2+m+\epsilon$，其中 n^2 是小于 x 的最大完全平方数，m 是非负整数，$0<\epsilon\leqslant 1$。则$\sqrt{x}$和$\sqrt{\lfloor x\rfloor}=\sqrt{n^2+m}$均在 n 和 $n+1$ 之间，因此式子两边均等于 n。

b)如果 x 是正整数，则两边相等。因此假定 $x=n^2-m-\epsilon$，其中 n^2 是大于 x 的最小完全平方数，m 是非负整数，$0<\epsilon\leqslant 1$。则$\sqrt{x}$和$\sqrt{\lceil x\rceil}=\sqrt{n^2-m}$均在 $n-1$ 和 n 之间，因此式子两边都等于 n。

69. a)定义域是 **Z**；伴域是 **R**；定义区域是非零整数集；f 的未定义区域是$\{0\}$；不是全函数。
b)定义域是 **Z**；伴域是 **R**；定义区域是 **Z**；f 的未定义区域是$\varnothing$；全函数。
c)定义域是 **Z×Z**；伴域是 **Q**；定义区域是 $\mathbf{Z}\times(\mathbf{Z}-\{0\})$；$f$ 的未定义区域是 $Z\times\{0\}$；不是全函数。
d)定义域是 **Z×Z**；伴域是 **Z**；定义区域是 **Z×Z**；f 的未定义区域是$\varnothing$；全函数。
e)定义域是 **Z×Z**；伴域是 **Z**；定义区域是$\{(m,n)\mid m>n\}$；f 的未定义区域是$\{(m,n)\mid m\leqslant n\}$；不是全函数。

71. 从公式显然可知，对于 $m+n$ 的固定值(例如 $m+n=x$)，函数的取值范围是从$(x-2)(x-1)/2+1$ 到$(x-2)(x-1)/2+(x-1)$，因为在这些条件下 m 可以取 1，2，3，…，$(x-1)$，并且在 $m+n$ 取固定值时公式中的第一项是固定的正整数。为了证明这个函数是一对一的且映上的，只须证明函数对于 $x+1$ 的取值范围，正好是函数对于不用的 x 的取值范围，即 $f(x-1,1)+1=f(1,x)$。我们有 $f(x-1,1)+1=\frac{(x-2)(x-1)}{2}+(x-1)+1=\frac{x^2-x+2}{2}=\frac{(x-1)x}{2}+1=f(1,x)$。

补充练习

1. a) $q\to p$　　b) $q\wedge p$　　c) $\neg q\vee\neg p$　　d) $q\leftrightarrow p$

3. a)除非 p 为假，即 p 为真，否则命题不会为假。如果 p 为真，q 为真，那么$\neg q\wedge(p\to q)$为假，蕴含为真。如果 p 为真而 q 为假，那么 $p\to q$ 为假，所以$\neg q\wedge(p\to q)$仍为假，蕴含还是为真。
b)除非 q 为假，否则命题不会为假。如果 q 为假，p 为真，那么$(p\vee q)\wedge\neg p$为假，蕴含为真。若 q 为假，p 为假，那么$(p\vee q)\wedge\neg p$也为假，蕴含仍为真。

5. $(p\wedge q\wedge r\wedge\neg s)\vee(p\wedge q\wedge\neg r\wedge s)\vee(p\wedge\neg q\wedge r\wedge s)\vee(\neg p\wedge q\wedge r\wedge s)$

7. 使用显而易见的字母，将这些语句翻译成符号，得到$\neg t\to\neg g$，$\neg g\to\neg q$，$r\to q$ 和$\neg t\wedge r$。假定这些语句一致。第四个语句表明$\neg t$必为真。因此对第一个语句，由假言推理可知$\neg g$为真，因此(从第二个语

句)可知$\neg q$为真。同样，第四个语句说明r必为真，对第三个语句，由假言推理q为真。这是一个矛盾：$q \wedge \neg q$。因此这些语句不一致。

9. Brenda

11. a)F　b)T　c)F　d)T　e)F　f)T

13. $\forall x \exists y \exists z(y \neq z \wedge \forall w(P(w, x) \leftrightarrow (w=y \vee w=z)))$

15. 假定$\exists x(P(x) \rightarrow Q(x))$为真。那么或者对某个$x_0$，$Q(x_0)$为真，这时$\forall x P(x) \rightarrow \exists x Q(x)$为真；或者有某个$x_0$，$P(x_0)$为假，这时$\forall x P(x) \rightarrow \exists x Q(x)$也为真。反过来，假定$\exists x(P(x) \rightarrow Q(x))$为假。这表明$\forall x(P(x) \wedge \neg Q(x))$为真，即$\forall x P(x)$为真，$\forall x(\neg Q(x))$也为真。后一命题等价于$\neg \exists x Q(x)$，因此$\forall x P(x) \rightarrow \exists x Q(x)$为假。

17. 不。

19. $\forall x \forall z \exists y T(x, y, z)$，其中$T(x, y, z)$为"学生$x$已选修$z$系的$y$课"，相应的论域是班上所有学生的集合，这所大学开设的所有课程的集合，以及数学学院所有系的集合。

21. a)$\overline{A}$　b)$A \cap B$　c)$A-B$　d)$\overline{A} \cap \overline{B}$　e)$A \oplus B$

23. 第二次使用存在量词消去时不能用字母c，因为它已经用过了。我们无法确保使得P为真的x与使得Q为真的x是相同的，因此不能将它们都称为c。

25. 令x为给定的。我们必须证明$\neg R(x) \rightarrow P(x)$，因此假设$\neg R(x)$为真。我们应用全称量词消去到第二个前提上，然后取逆否并应用德摩根律得到$\neg R(x) \rightarrow (P(x) \vee \neg Q(x))$。由假言推理，这给出$P(x) \vee \neg Q(x)$。由全称量词消去，第一个前提给出$P(x) \vee Q(x)$。通过消解结合这些，我们有$P(x)$，因此得证。

27. 在此给出间接证明，证明若$\sqrt{x}$是有理数，则x是有理数，假设$x \geqslant 0$。假定$\sqrt{x}=p/q$是有理数，$q \neq 0$。则$x=(\sqrt{x})^2=p^2/q^2$也是有理数(q^2也是非零的)。

29. 通过令$m=10^{500}+1$，可以给出构造性证明。则$m^2=(10^{500}+1)^2>(10^{500})^2=10^{1000}$。

31. 23不能写为8个立方体的和。

33. 223不能写为36个5次幂的和。

35. 是。

37. a)$A \cap \overline{A}=\{x \mid x \in A \wedge x \notin A\}=\varnothing$

b)$A \cup \overline{A}=\{x \mid x \in A \vee x \notin A\}=U$

39. $A-(A-B)=A-(A \cap \overline{B})=A \cap (\overline{A \cap \overline{B}})=A \cap (\overline{A} \cup B)$

$=(A \cap \overline{A}) \cup (A \cap B)=\varnothing \cup (A \cap B)=A \cap B$

41. 令$A=\{1\}$，$B=\varnothing$，$C=\{1\}$，于是$(A-B)-C=\varnothing$，但$A-(B-C)=\{1\}$。

43. 不等。例如，令$A=B=\{a, b\}$，$C=\varnothing$，$D=\{a\}$。于是$(A-B)-(C-D)=\varnothing-\varnothing=\varnothing$，但$(A-C)-(B-D)=\{a, b\}-\{b\}=\{a\}$。

45. a)$|\varnothing| \leqslant |A \cap B| \leqslant |A| \leqslant |A \cup B| \leqslant |U|$

b)$|\varnothing| \leqslant |A-B| \leqslant |A \oplus B| \leqslant |A \cup B| \leqslant |A|+|B|$

47. a)f是，g不是。

b)f是，g不是。

c)f有逆：$f^{-1}(a)=3$，$f^{-1}(b)=4$，$f^{-1}(c)=2$，$f^{-1}(d)=1$；g没有逆。

49. 令$f(a)=f(b)=1$，$f(c)=f(d)=2$，$S=\{a, c\}$，$T=\{b, d\}$。于是$f(S \cap T)=f(\varnothing)=\varnothing$，但$f(S) \cap f(T)=\{1, 2\} \cap \{1, 2\}=\{1, 2\}$。

51. 此等式为真当且仅当x和y的小数部分的和小于1。

第2章

2.1节

1. $max:=1$，$i:=2$，$max:=8$，$i:=3$，$max:=12$，$i:=4$，$i:=5$，$i:=6$，$i:=7$，$max:=14$，$i:=8$，

$i:=9$，$i:=10$，$i:=11$

3. **procedure** *sum*(a_1，…，a_n：整数)

   $sum:=a_1$

   for $i:=2$ **to** n

   　$sum:=sum+a_i$

   {*sum* 为所求之值}

5. **procedure** *duplicates*(a_1，a_2，…，a_n：按非增顺序排列的整数)

   $k:=0${这个变量计算重复出现数的个数}

   $j:=2$

   while $j\leqslant n$

   begin

   　if $a_j=a_{j-1}$ **then**

   　begin

   　　$k:=k+1$

   　　$c_k:=a_j$

   　　while($j\leqslant n$ 且 $a_j=c_k$)

   　　　$j:=j+1$

   　end

   $j:j+1$

   end{c_1，c_2，…，c_k 为所求之表}

7. **procedure** *last even location*(a_1，a_2，…，a_n：整数)

   $k:=0$

   for$i:=1$ **to** n

   　if a_i 是偶数 **then** $k:=i$

   end{k 为所求的位置，若没有偶数则 k 为 0}

9. **procedure** *palindrome check*($a_1a_2\cdots a_n$：字符串)

   $answer:=$**true**

   for $i:=1$ **to** $\lfloor n/2\rfloor$

   　if $a_i\neq a_{n+1-i}$ **then** $answer:=$**false**

   end {*answer* 为真当且仅当该串为回文}

11. **procedure** *interchange*(x，y：实数)

    $z:=x$

    $x:=y$

    $y:=z$

    至少需要 3 个赋值语句。

13. 线性搜索：$i:=1$，$i:=2$，$i:=3$，$i:=4$，$i:=5$，$i:=6$，$i:=7$，$location:=7$；二分搜索：$i:=1$，$j:=8$，$m:=4$，$i:=5$，$m:=6$，$i:=7$，$m:=7$，$j:=7$，$location:=7$

15. **procedure** *insert*(x，a_1，a_2，…，a_n：按非增顺序排列的整数)

    {这个表的顺序是 $a_1\leqslant a_2\leqslant\cdots\leqslant a_n$}

    $a_{n+1}:=x+1$

    $i:=1$

    while $x>a_i$

    　$i:=i+1$

    for $j:=0$ **to** $n-i$

    　$a_{n-j+1}:=a_{n-j}$

```
    a_i := x
    {x 已被插入到正确位置}
17. procedure first largest(a_1, …, a_n: 整数)
    max := a_1
    location := 1
    for i := 2 to n
    begin
      if max < a_i then
      begin
        max := a_i
        location := i
      end
    end
19. procedure mean-median-max-min(a, b, c: 整数)
    mean := (a+b+c)/3
    {分别处理 a, b, c 的 6 种不同顺序}
    if a > b then
    begin
        if b > c then
        median := b; max := a; min := c
      end
      …
    (算法的其余部分是类似的)
21. procedure first-three(a_1, a_2, …, a_n: 整数)
    if a_1 > a_2 then 交换 a_1 与 a_2
      if a_2 > a_3 then 交换 a_2 与 a_3
      if a_1 > a_2 then 交换 a_1 与 a_2
23. procedure onto(f: 从 A 到 B 的函数, 其中 A={a_1, …, a_n}, B={b_1, …, b_m}, a_1, …, a_n, b_1, …,
    b_m 是整数)
    for i := 1 to m
      hit(b_i) := 0
      count := 0
    for j := 1 to n
      if hit(f(a_j)) = 0 then
      begin
          hit(f(a_j)) := 1
          count := count+1
      end
    if count = m then onto := true
    else onto := false
25. procedure ones(a: 位串, a=a_1a_2…a_n)
    ones := 0
    for i := 1 to n
    begin
        if a_i := 1 then
          ones := ones+1
```



```
    end{ones 是位串 a 中 1 的个数}
27. procedure ternary search(s: 整数, a1, a2, …, an: 递增的整数)
   i := 1
   j := n
   while i < j − 1
   begin
       l = ⌊(i + j)/3⌋
       u = ⌊2(i + j)/3⌋
       if x > au then i := u + 1
       else if x > al then
       begin
           i := l + 1
           j := u
       end
        else j := l
     end
   if x = ai then location := i
   else if x = aj then location := j
   else location := 0
   {location 是等于 x 那一项的下标(若没有找到则为 0)}
29. procedure find a mode(a1, a2, …, an: 非递减的整数)
   modecount := 0
   i := 1 while i ≤ n
   begin
      value := ai
      count := 1
      while i ≤ n 且 ai = value
      begin
          count := count + 1
          i := i + 1
        end
      if count > modecount then
      begin
          modecount := count
          mode := value
      end
     end
   {mode 是第一个最常出现的值}
31. procedure find duplicate(a1, a2, …, an: 整数)
   location := 0
   i := 2
   while i ≤ n 且 location = 0
   begin
     j := 1
```



```
    while j<i 且 location =0
      if a_i=a_j then location :=i
      else j :=j+1
      i :=i+1
    end
  {location 是序列中第一个重复项的下标}
```

33. **procedure** *find decrease*(a_1, a_2, …, a_n : 正整数)

```
  location :=0
  i :=2
  while i≤n 且 location=0
    if a_i<a_{i-1} then location :=i
    else i :=i+1
  {location 是第一个比前一项小的值的下标}
```

35. 第一遍结束时：1，3，5，4，7；第二遍结束时：1，3，4，5，7；第三遍结束时；1，3，4，5，7；第四遍结束时：1，3，4，5，7。

37. **procedure** *better bubblesort*(a_1, …, a_n: 整数)

```
  i :=1; done :=false
  while(i<n 且 done=false)
  begin
    done :=true
    for j :=1 to n-i
      if a_j>a_{j+1} then
      begin
        交换 a_j 与 a_{j+1}
        done :=false
      end
    i :=i+1
  end{a_1, …, a_n 按递增顺序排列}
```

39. 第一、二、三遍结束时：1，3，5，7，4；第四遍结束时：1，3，4，5，7。

41. a)1，5，4，3，2；1，2，4，3，5；1，2，3，4，5；1，2，3，4，5

b)1，4，3，2，5；1，2，3，4，5；1，2，3，4，5；1，2，3，4，5

c)1，2，3，4，5；1，2，3，4，5；1，2，3，4，5；1，2，3，4，5

43. 运行本节中给出的算法 2，但是把 $x\neq a_i$ 换成 $x<a_i$，并且把 **else** 语句换成 **else** *location* :=$n+1$。

45. $2+3+4+\cdots+n=(n^2+n-2)/2$

47. 在表 3 中找 2 的位置(1 次比较)，把 2 插入 3 的前面，现在表看起来是 2，3，4，5，1，6。找 4 的位置(先与 2 再与 3 比较)，插入 4，得到 2，3，4，5，1，6。找 5 的位置(先与 3 再与 4 比较)，插入 5，得到 2，3，4，5，1，6。找 1 的位置(先与 3 和 2 比较，再与 2 比较)，插入 1，得到 1，2，3，4，5，6。找 6 的位置(先与 3 和 4 比较，再与 5 比较)，插入 6，得到最终的答案 1，2，3，4，5，6。

49. **procedure** *binary insertion sort*(a_1, a_2, …, a_n: 实数，$n\geq 2$)

```
  for j :=2 to n
  begin
    {对位置 i 二分检索}
    left :=1
    right :=j-1
    while left<right
```

```
    begin
    middle := ⌊(left+right)/2⌋
    if a_j > a_middle then left := middle+1
    else right := middle
  end
  if a_j < a_left then i := left else i := left+1
  {把 a_j 插入位置 i，把 a_i 到 a_{j-1} 向后移动}
  m := a_j
  for k := 0 to j-i-1
    a_{j-k} := a_{j-k-1}
  a_i := m
  end{a_1, a_2, …, a_n 已被排序}
```

51. 练习 50 描述的变种更好。

53. a)2 个 25 美分，1 个 1 美分　　b)2 个 25 美分，1 个 10 美分，1 个 5 美分，4 个 1 美分
    c)3 个 25 美分，1 个 1 美分　　d)2 个 25 美分，1 个 10 美分

55. 贪心算法对 a)、c)、d)使用最少的硬币。
    a)2 个 25 美分，1 个 1 美分　　b)2 个 25 美分，1 个 10 美分，9 个 1 美分
    c)3 个 25 美分，1 个 1 美分　　d)2 个 25 美分，1 个 10 美分

57. a)变量 f 给出最后选择的那个演讲的结束时间，f 起初等于报告厅可以开始使用的时间。把演讲按照结束时间的递增顺序排序，从表的顶端开始。在算法每个阶段，从表中剩余演讲向下，找第一个开始时间不早于 f 的演讲。安排那个报告，更新 f 以记录那个演讲的结束时间。
    b)9：00～9：45 演讲，9：50～10：15 演讲，10：15～10：45 演讲，11：00～11：15 演讲。

2.2 节

1. 对 C 和 k 的选择不唯一。
   a)$C=1$，$k=10$　　b)$C=4$，$k=7$　　c)不是
   d)$C=5$，$k=1$　　e)$C=1$，$k=0$　　f)$C=1$，$k=2$

3. 对所有 $x>9$，$x^4+9x^3+4x+7\leqslant 4x^4$；见证 $C=4$，$k=9$。

5. 对所有 $x>1$，$(x^2+1)-(x+1)=x-1+2/(x+1)<x$；见证 $C=1$，$k=1$。

7. 对 C 和 k 的选择不唯一。
   a)$n=3$，$C=3$，$k=1$　　b)$n=3$，$C=4$，$k=1$
   c)$n=1$，$C=2$，$k=1$　　d)$n=0$，$C=2$，$k=1$

9. 对所有 $x>17$，$x^2+4x+17\leqslant 3x^3$，所以 $x^2+4x+17$ 是 $O(x^3)$，见证 $C=3$，$k=17$。可是若 x^3 是 $O(x^2+4x+17)$，那么对某个 C 和所有足够大的 x，$x^3\leqslant C(x^2+4x+17)\leqslant 3Cx^2$，这表明对所有足够大的 x 有 $x<C$，这是不可能的，所以 x^3 不是 $O(x^2+4x+17)$。

11. 对 $x>1$，$3x^4+1\leqslant 4x^4=8(x^4/2)$，所以 $3x^4+1$ 是 $O(x^4/2)$，见证是 $C=4$，$k=8$。另外，对 $x>0$，$x^4/2\leqslant 3x^4+1$，所以 $x^4/2$ 是 $O(3x^4+1)$，见证是 $C=1$，$k=0$。

13. 因为对所有 $n>0$，$2^n\leqslant 3^n$，所以 2^n 是 $O(3^n)$。但如果 3^n 也是 $O(2^n)$，则对某个 C，对所有足够大的 n，$3^n\leqslant C\cdot 2^n$ 成立。这说明对所有足够大的 n 有 $C\geqslant(3/2)^n$，但这是不可能的。所有 3^n 不是 $O(2^n)$。

15. 即所有这样的函数，存在实数 k 和 C，对 $x>k$ 有 $|f(x)|\leqslant C$。这些函数是对足够大的 x 有上界的函数。

17. 存在常数 C_1，C_2，k_1，k_2，使得对所有 $x>k_1$ 有 $|f(x)|\leqslant C_1|g(x)|$，对所有 $x>k_2$ 有 $|g(x)|\leqslant C_2|h(x)|$。因此对 $x>\max(k_1, k_2)$有 $|f(x)|\leqslant C_1|g(x)|\leqslant C_1C_2|h(x)|$。这说明 $f(x)$是 $O(h(x))$。

19. a)$O(n^3)$ b)$O(n^5)$ c)$O(n^3 \cdot n!)$

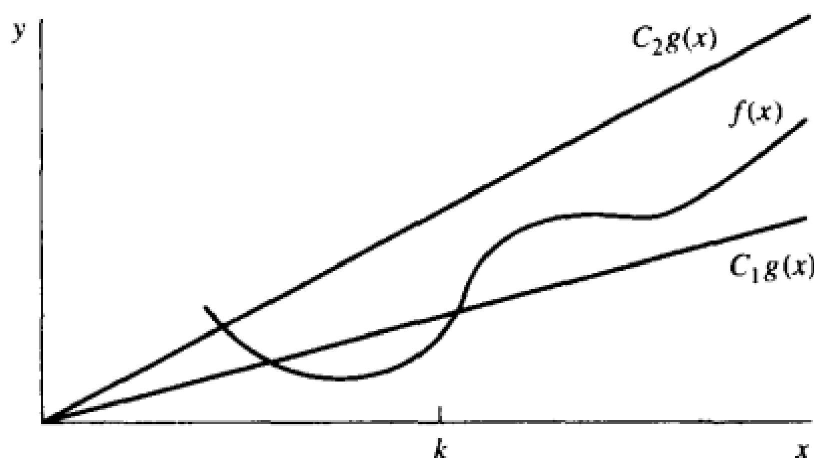
21. a)$O(n^2 \log n)$ b)$O(n^2(\log n)^2)$ c)$O(n^{2^n})$

23. a)既非$\Theta(x^2)$，也非$\Omega(x^2)$。 b)$\Theta(x^2)$和$\Omega(x^2)$。
c)既非$\Theta(x^2)$，也非$\Omega(x^2)$。 d)$\Omega(x^2)$，但非$\Theta(x^2)$。
e)$\Omega(x^2)$，但非$\Theta(x^2)$。 f)$\Omega(x^2)$和$\Theta(x^2)$。

25. 如果$f(x)$是$\Theta(g(x))$，则有常数C_1，C_2使$C_1|g(x)| \leqslant |f(x)| \leqslant C_2|g(x)|$。于是对$x>k$有$|f(x)| \leqslant C_2|g(x)|$，$|g(x)| \leqslant (1/C_1)|f(x)|$。因此$f(x)$是$O(g(x))$，$g(x)$是$O(f(x))$。反之，假设$f(x)$是$O(g(x))$，$g(x)$是$O(f(x))$。则存在常数$C_1$，$C_2$，$k_1$，$k_2$，使得对$x>k_1$有$|f(x)| \leqslant C_1|g(x)|$，对$x>k_2$有$|g(x)| \leqslant C_2|f(x)|$。可以假设$C_2>0$(总可以把$C_2$取得更大)，于是对$x>\max(k_1, k_2)$有$(1/c_2)|g(x)| \leqslant |f(x)| \leqslant C_1|g(x)|$。所以$f(x)$是$\Theta(g(x))$。

27. 如果$f(x)$是$\Theta(g(x))$，则$f(x)$既是$O(g(x))$，又是$\Omega(g(x))$。因此存在正常数C_1，C_2，k_1，k_2，使得对所有$x>k_2$有$|f(x)| \leqslant C_2|g(x)|$，对所有$x>k_1$有$|f(x)| \geqslant C_1|g(x)|$。于是只要$x>k$，其中$k=\max(k_1k_2)$，就有$C_1|g(x)| \leqslant |f(x)| \leqslant C_2|g(x)|$。反之，如果有正常数$C_1$，$C_2$和$k$，使得对所有$x>k$有$C_1|g(x)| \leqslant |f(x)| \leqslant C_2|g(x)|$，则令$k_1=k_2=k$，可知$f(x)$既是$O(g(x))$，又是$\Theta(g(x))$。

29.



31. 如果$f(x)$是$\Theta(1)$，那么$|f(x)|$限于正常数C_1和C_2之间。换句话说，$f(x)$不可能增大到超过一个固定的界，也不能小于这个界的负数，而且必定不会比某个固定的界更接近于0。

33. 由于$f(x)$是$O(g(x))$，所以一定有常数C和l，使得对所有x，$x>l$，有$|f(x)| \leqslant C|g(x)|$。于是对所有$x>1$有$|f^k(x)| \leqslant C^k|g^k(x)|$。所以取常数为$C^k$，则$f^k(x)$是$O(g^k(x))$。

35. 由于$f(x)$和$g(x)$均是增长的且有界，故可假设对足够大的x，有$f(x) \geqslant 1$及$g(x) \geqslant 1$。存在常数C和k，使得对于$x>k$有$f(x) \leqslant Cg(x)$。这说明对足够大的x，有$\log f(x) \leqslant \log C + \log g(x) < 2\log g(x)$。因此$\log f(x)$是$O(\log g(x))$。

37. 根据定义，有常数C_1，C_1'，C_2，C_2'，k_1，k_1'，k_2，k_2'，使得对所有$x>k_1$有$f_1(x) \geqslant C_1|g(x)|$，对所有$x>k_1'$有$f_1(x) \leqslant C_1'|g(x)|$；对所有$x>k_2$有$f_2(x) \geqslant C_2|g(x)|$，对所有$x>k_2'$有$f_2(x) \leqslant C_2'|g(x)|$。把第一个和第三个不等式相加，说明对所有$x>k$，$k=\max(k_1, k_2)$有$f_1(x)+f_2(x) \geqslant (C_1+C_2)|g(x)|$；把第二个和第四个不等式相加，说明对所有$x>k'$，其中，$k'=\max(k_1', k_2')$，有$f_1(x)+f_2(x) \leqslant (C_1'+C_2')|g(x)|$。因此$f_1(x)+f_2(x)$是$\Theta(g(x))$。如果$f_1$和$f_2$能取负值，则这一结论就不一定对了。

39. 结论不成立。令$f_1(x)=x^2+2x$，$f_2(x)=x^2+x$，$g(x)=x^2$。则$f_1(x)$和$f_2(x)$都是$\Theta(g(x))$，但$(f_1-f_2)(x)$不是。

41. 取$f(n)$是这样的函数，如果n是奇正数，则$f(n)=n$；如果n是偶正数，则$f(n)=1$；$g(n)$则相反，如果n是奇正数，则$g(n)=1$；如果n是偶正数，则$g(n)=n$。

43. 存在正常数C_1，C_1'，C_2，C_2'，k_1，k_1'，k_2，k_2'，使得对所有$x>k_1$有$|f_1(x)| \geqslant C_1|g_1(x)|$，对所有$x>k_1'$有$|(f_1(x))| \leqslant C_1'|g_1(x)|$，对所有$x>k_2$有$|f_2(x)| > C_2|g_2(x)|$，而对所有$x>k_2'$有$|f_2(x)| \leqslant C_2'|g_2(x)|$。由于$f_2$和$g_2$总不为0，最后两个不等式可以重写为，对所有$x>k_2$有

$|1/f_2(x)| \leqslant (1/C_2)|1/g_2(x)|$，对所有 $x>k_2'$ 有 $|1/f_2(x)| \geqslant (1/C_2')|1/g_2(x)|$。把第一个不等式和重写过的第四个不等式相乘，说明对所有 $x>\max(k_1, k_2')$，有 $|f_1(x)/f_2(x)| \geqslant (C_1/C_2')|g_1(x)/g_2(x)|$；把第二个不等式和重写过的第三个不等式相乘，说明对所有 $x>\max(k_2, k_1')$，有 $|f_1(x)/f_2(x)| \geqslant (C_1'/C_2)|g_1(x)/g_2(x)|$。因此 f_1/f_2 是 $\Theta(g_1/g_2)$。

45. 存在正常数 C_1，C_2，k_1，k_2，k_1'，k_2'，使得对所有 $x>k_1$，$y>k_2$，有 $|f(x, y)| \leqslant C_1|g(x, y)|$，对所有 $x>k_1'$，$y>k_2'$，有 $|f(x, y)| \geqslant C_2|g(x, y)|$。

47. 对 $x>1$ 和 $y>1$，有 $(x^2+xy+x\log y)^3<(3x^2y)^3=27x^6y^3$，因为 $x^2<x^2y$，$xy<x^2y$ 及 $x\log y<x^2y$。因此 $(x^2+xy+x\log y)^3$ 是 $O(x^6y^3)$。

49. 对所有正实数 x 和 y，有 $\lfloor xy \rfloor \leqslant xy$。因此取 $C=1$，$k_1=k_2=0$，由定义知 $\lfloor xy \rfloor$ 是 $O(xy)$。

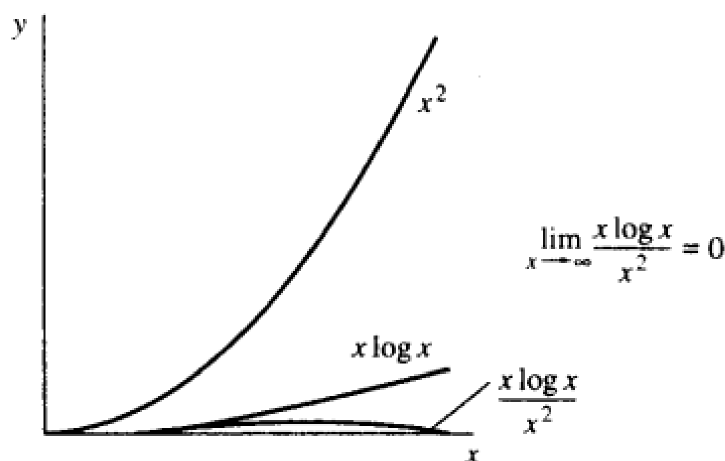
51. a) $\lim\limits_{x\to\infty} x^2/x^3=\lim\limits_{x\to\infty} 1/x=0$

b) $\lim\limits_{x\to\infty}\dfrac{x\log x}{x^2}=\lim\limits_{x\to\infty}\dfrac{\log x}{x}=\lim\limits_{x\to\infty}\dfrac{1}{x\ln 2}=0$ （用洛必达法则）

c) $\lim\limits_{x\to\infty}\dfrac{x^2}{2^x}=\lim\limits_{x\to\infty}\dfrac{2x}{2^x\ln 2}=\lim\limits_{x\to\infty}\dfrac{2}{2^x(\ln 2)^2}=0$ （用洛必达法则）

d) $\lim\limits_{x\to\infty}\dfrac{x^2+x+1}{x^2}=\lim\limits_{x\to\infty}\left(1+\dfrac{1}{x}+\dfrac{1}{x^2}\right)=1\neq 0$

53.



55. 不是。取 $f(x)=1/x^2$ 和 $g(x)=1/x$。

57. a) 由于 $\lim\limits_{x\to\infty} f(x)/g(x)=0$，对足够大的 x 有 $|f(x)|/|g(x)|<1$。因此对某个常数 k，对 $x>k$ 有 $|f(x)|<|g(x)|$。于是 $f(x)$ 是 $O(g(x))$。

b) 令 $f(x)=g(x)=x$。则 $f(x)$ 是 $O(g(x))$，但 $f(x)$ 不是 $O(g(x))$，因为 $f(x)/g(x)=1$。

59. 由于 $f_2(x)$ 是 $O(g(x))$，由练习 57(a) 得 $f_2(x)$ 是 $O(g(x))$。由推论 1 有 $f_1(x)+f_2(x)$ 是 $O(g(x))$。

61. 容易证明：对 $i=0, 1, \cdots, n-1$，有 $(n-i)(i+1)\geqslant n$。因此 $(n!)^2=(n\cdot 1)((n-1)\cdot 2)((n-2)\cdot 3)\cdot\cdots\cdot(2\cdot(n-1))(1\cdot n)\geqslant n^n$，因此 $2\log n! \geqslant n\log n$。

63. a) 渐近　　b) 非渐近　　c) 渐近

d) 渐近　　e) 非渐近　　f) 非渐近

g) 非渐近

2.3 节

1. $2n-1$

3. 线性检索

5. $O(n)$

7. a) $power:=1$，$y:=1$，$i:=1$，$power:=2$，$y:=3$，$i:=2$，$power:=4$，$y:=15$。

b) $2n$ 次乘法和 n 次加法

9. a) $2^{10^9}\approx 10^{3\times 10^8}$　　b) 10^9　　c) 3.96×10^7

d)3.16×10^4　　e)29　　f)12

11. a)36 年　　b)13 天　　c)19 分钟

13. 平均比较次数为$(3n+4)/2$

15. $O(\log n)$

17. $O(n)$

19. $O(n^2)$

21. $O(n)$

23. $O(n)$

25. $O(\log n)$次比较；$O(n^2)$次交换

27. a)加倍　　b)加 1

2.4 节

1. a)能　　b)不能　　c)能　　d)不能

3. 假设 $a\mid b$。则存在整数 k，使 $ak=b$。由于 $a(ck)=bc$ 所以 $a\mid bc$。

5. 如果 $a\mid b$ 且 $b\mid a$，则存在整数 c 和 d，使 $b=ac$ 且 $a=bd$。因此 $a=acd$。由于 $a\neq0$，所以$cd=1$。因此，要么 $c=d=1$，要么 $c=d=-1$。因此要么 $a=b$，要么 $a=-b$。

7. 由于 $ac\mid bc$，必有整数 k，使 $ack=bc$，从而 $ak=b$，于是 $a\mid b$。

9. a)2，5　　b)−11，10　　c)34，7　　d)77，0
   e)0，0　　f)0，3　　g)−1，2
   h)4，0

11. a)$2^3\cdot11$　　b)$2\cdot3^2\cdot7$　　c)3^6　　d)$7\cdot11\cdot13$
   e)$11\cdot101$　　f)$2\cdot3^3\cdot5\cdot7\cdot13\cdot37$

13. $2^8\cdot3^4\cdot5^2\cdot7$

15. 假定 $\log_2 3=a/b$，其中 a，$b\in\mathbf{Z}^+$ 且 $b\neq0$。于是 $2^{a/b}=3$，从而 $2^a=3^b$。这不符合算术基本定理。因此 $\log_2 3$ 是无理数。

17. 1，7，11，13，17，19，23，29

19. a)是　　b)不是　　c)是　　d)是

21. 如果 $a \bmod m=b \bmod m$，则 a 和 b 除以 m 时余数相同。于是 $a=q_1m+r$，$b=q_2m+r$，其中$0\leqslant r<m$。所以 $a-b=(q_1-q_2)m$，故 $m\mid(a-b)$。所以 $a\equiv b \pmod m$。

23. 假定 n 不是素数，于是 $n=ab$，其中 a 和 b 为大于 1 的整数。由于 $a>1$，根据提示，2^a-1 是 2^n-1 的一个大于 1 的因子，2^n-1 的另一个因子也大于 1，所以 2^n-1 不是素数。

25. a)2　　b)4　　c)12

27. $\phi(p^k)=p^k-p^{k-1}$

29. a)$3^5\cdot5^3$　　b)1　　c)23^{17}　　d)$41\cdot43\cdot53$
   e)1　　f)1111

31. a)$2^{11}\cdot3^7\cdot5^9\cdot7^3$　　b)$2^9\cdot3^7\cdot5^5\cdot7^3\cdot11\cdot13\cdot17$　　c)23^{31}
   d)$41\cdot43\cdot53$　　e)$2^{12}\cdot3^{13}\cdot5^{17}\cdot7^{21}$　　f)无定义

33. 存在某个 b 使$(b-1)k<n\leqslant bk$。因此$(b-1)k\leqslant n-1<bk$。两边除以 k 得到 $b-1<n/k\leqslant b$ 和 $b-1\leqslant(n-1)/k<b$。因此$\lceil n/k\rceil=b$ 且$\lfloor(n-1)/k\rfloor=b-1$。

35. 若 $x \bmod m\leqslant\lceil m/2\rceil$，则 $x \bmod m$；若 $x \bmod m>\lceil m/2\rceil$，则$(x \bmod m)-m$。

37. a)1　　b)2　　c)3　　d)9

39. a)否　　b)否　　c)是　　d)否

41. 由于 $\min(x, y)+\max(x, y)=x+y$，所以 p_i 在 $\gcd(a, b)\cdot\mathrm{lcm}(a, b)$的素因子分解中的指数是 p_i 在 a 和 b 的素因子分解中指数之和。

43. 令 $m=tn$。由于 $a\equiv b(\bmod m)$，存在整数 s 使 $a=b+sm$。因此 $a=b+(st)n$，所以 $a\equiv b(\bmod n)$。
45. 令 $m=c=2$，$a=0$ 和 $b=1$。于是 $0=ac\equiv bc=2(\bmod 2)$，但 $0=a\neq b=1(\bmod 2)$。
47. 由于 $a\equiv b(\bmod m)$，存在整数 s 使得 $a=b+sm$，所以 $a-b=sm$。于是 $k\geqslant 2$ 时，$a^k-b^k=(a-b)(a^{k-1}+a^{k-2}b+\cdots+ab^{k-2}+b^{k-1})$也是 m 的倍数。于是 $a^k\equiv b^k(\bmod m)$。
49. a)7，19，7，7，18，0
    b)取下一可用空间 mod 31。
51. 2，6，7，10，8，2，6，7，10，8，…
53. a)GR QRW SDVV JR　b) QB ABG CNFF TB　c) QX UXM AHJJ ZX
55. 0
57. 本书的 ISBN 检验位是有效的，因为 $1\cdot 0+2\cdot 0+3\cdot 7+4\cdot 2+5\cdot 4+6\cdot 2+7\cdot 4+8\cdot 3+9\cdot 4+10\cdot 6\equiv 0(\bmod 11)$。
59. a)如果 n 为素数，$a_n=1$；否则 $a_n=0$。
    b)a_n 是 n 的最小素因子，且 $a_1=1$。
    c)a_n 是 n 的正因子的个数。
    d)如果 n 没有大于 1 的完全平方正因子，$a_n=1$，否则 $a_n=0$。
    e)a_n 是小于或等于 n 的最大素数。
    f)a_n 是前 $n-1$ 个素数之积。

2.5 节

1. a)1110 0111　b)1 0001 1011 0100　c)10111 1101 0110 1100
3. a)31　b)513　c)341　d)26 896
5. a)1000 0000 1110
   b)1 0011 0101 1010 1011
   c)1010 1011 1011 1010
   d)1101 1110 1111 1010 1100 1110 1101
7. 1010 1011 1100 1101 1110 1111
9. $(B7B)_{16}$
11. 必要时在开头加 3 个零，把二进制展开式写成$(\cdots b_{23}b_{22}b_{21}b_{20}b_{13}b_{12}b_{11}b_{10}b_{03}b_{02}b_{01}b_{00})_2$。这个数字的值是 $b_{00}+2b_{01}+4b_{02}+8b_{03}+2^4b_{10}+2^5b_{11}+2^6b_{12}+2^7b_{13}+2^8b_{20}+2^9b_{21}+2^{10}b_{22}+2^{11}b_{23}+\cdots$，这个值可以写成 $b_{00}+2b_{01}+4b_{02}+8b_{03}+(b_{10}+2b_{11}+4b_{12}+8b_{13})\cdot 2^4+(b_{20}+2b_{21}+4b_{22}+8b_{23})\cdot 2^8+\cdots$。现在把$(b_{i3}b_{i2}b_{i1}b_{i0})_2$ 转换为十六进制 h_i。所以这个数是 $h_0+h_1\cdot 2^4+h_2\cdot 2^8+\cdots=h_0+h_1\cdot 16+h_2\cdot 16^2+\cdots$，即十六进制展开式$(\cdots h_2h_1h_0)_{16}$。
13. 把二进制数字 3 个分为一组，必要时加 2 个零，把每组二进制数字转换成 1 个八进制数字。
15. $(1110111001010111010001)_2$，$(1273)_8$
17. 用练习 14 把给定的八进制数转换为二进制，再用例 6 从二进制转化成十六进制。
19. 27
21. a)6　b)3　c)11
    d)3　e)40　f)12
23. 8
25. 整数的二进制展开是唯一的这种和。
27. 令 $a=(a_{n-1}a_{n-2}\cdots a_1a_0)_{10}$，那么 $a=10^{n-1}a_{n-1}+10^{n-2}a_{n-2}+\cdots+10a_1+a_0\equiv a_{n-1}+a_{n-2}+\cdots+a_1+a_0(\bmod 3)$，这是因为对所有非负整数 j，$10^j\equiv 1(\bmod 3)$。于是 $3\mid a$ 当且仅当 3 能除尽 a 的十进数字之和。
29. 令 $a=(a_{n-1}a_{n-2}\cdots a_1a_0)_2$，那么 $a=a_0+2a_1+2^2a_2+\cdots+2^{n-1}a_{n-1}\equiv a_0-a_1+a_2-a_3+\cdots\pm a_{n-1}(\bmod 3)$。于是 a 能被 3 整除当且仅当偶数位的二进数字之和减去奇数位的二进数字之和能被 3 整除。

31. a) -6 b) 13 c) -14 d) 0

33. 和的 1 补表示可以这样计算：两个整数的 1 补相加，不过首位的进位要用做和的末位的进位。

35. 若 $m \geqslant 0$，那么 m 的 1 补展开的首位 a_{n-1} 为 0，公式为 $m=\sum_{i=0}^{n-2} a_i 2^i$。这一公式正确，因为右边正是 m 的二进展开。若 m 是负的，那么 m 的 1 补展开的首位是 1。其他的 $n-1$ 位可以用 $111\cdots1$（共 $n-1$ 个 1）减去 $-m$ 得到，因为 1 减去一个字位就是该字位的补。于是位串 $a_{n-2}a_{n-3}\cdots a_1a_0$ 恰是 $(2^{n-1}-1)-(-m)$ 的二进展开。解方程 $(2^{n-1}-1)-(-m)=\sum_{i=0}^{n-2} a_i 2^i$ 求 m 即可得想证明的等式，因为 $a_{n-1}=1$。

37. a) -7 b) 13 c) -15 d) -1

39. 要得到两个整数之和的 2 补表示，把它们的 2 补表示相加（就像加两个二进整数一样），忽略最左边一列的进位。不过若出现溢出，则结果就不正确了。发生溢出的情况是这两项的 2 补表示的最左位数字相同，而答案的最左位数字与它们不同。

41. 若 $m \geqslant 0$，那么首位 a_{n-1} 为 0，公式为 $m=\sum_{i=0}^{n-2} a_i 2^i$。由于右边正是 m 的二进展开，所以公式正确。若 $m<0$，它的 2 补展开以 1 为首位，其余 $n-1$ 位是 $2^{n-1}-(-m)$ 的二进展开。这表明 $(2^{n-1})-(-m)=\sum_{i=0}^{n-2} a_i 2^i$。解此方程求 m 即得想求的等式，因为 $a_{n-1}=1$。

43. $4n$

45.
```
procedure Cantor(x: 正整数)
n := 1; f := 1
while (n+1)·f ≤ x
begin
    n := n+1
    f := f·n
  end
y := x
while n > 0
begin
    a_n := ⌊y/f⌋
    y := y − a_n·f
    f := f/n
    n := n−1
  end{x = a_n n! + a_{n−1}(n−1)! + ⋯ + a_1 1!}
```

47. 第一步：$c=0$，$d=0$，$s_0=1$；第二步：$c=0$，$d=1$，$s_1=0$；第三步：$c=1$，$d=1$，$s_2=0$；第四步：$c=1$，$d=1$，$s_3=0$；第五步：$c=1$，$d=1$，$s_4=1$；第六步：$c=1$，$s_5=1$

49.
```
procedure subtract(a, b: 正整数, a>b, a=(a_{n−1}a_{n−2}⋯a_1a_0)_2, b=(b_{n−1}b_{n−2}⋯b_1b_0)_2)
B := 0{B是借位}
for j := 0 to n−1
begin
    if a_j ≥ b_j + B then
    begin
      s_j := a_j − b_j − B
      B := 0
    end
    else
```

```
        begin
          s_j := a_j + 2 - b_j - B
          B := 1
        end
      end{(s_{n-1} s_{n-2} … s_1 s_0)_2 是差}
```

51. **procedure** *compare*(a，b：正整数，且 $a=(a_na_{n-1}\cdots a_1a_0)_2$，$b=(b_nb_{n-1}\cdots b_1b_0)_2$)

```
k := n
while a_k = b_k 且 k > 0
    k := k - 1
if a_k = b_k then print "a 等于 b"
if a_k > b_k then print "a 大于 b"
if a_k < b_k then print "a 小于 b"
```

53. $O(\log n)$

55. 算法中仅有的消耗时间部分是 while 循环，共迭代 q 次。循环内做的工作是把两个不超过 $|a|$ 的数相减，这两个数各有 $\log|a|$ 位。从例 8 就得出结论。

2.6 节

1. a) $1=(-1)\cdot 10+1\cdot 11$ b) $1=21\cdot 21+(-10)\cdot 44$
   c) $12=(-1)\cdot 36+48$ d) $1=13\cdot 55+(-21)\cdot 34$
   e) $3=11\cdot 213+(-20)\cdot 117$ f) $223=1\cdot 0+1\cdot 223$
   g) $1=37\cdot 2347+(-706)\cdot 123$ h) $2=1128\cdot 3454+(-835)\cdot 4666$
   i) $1=2468\cdot 9999+(-2221)\cdot 11111$

3. $15\cdot 7=105\equiv 1\pmod{26}$

5. 7

7. 52

9. 假定 b 和 c 都是 a 模 m 的逆，那么 $ba\equiv 1\pmod m$，$ca\equiv 1\pmod m$。因此 $ba\equiv ca\pmod m$。由于 $\gcd(a, m)=1$，由定理 2 知 $b\equiv c\pmod m$。

11. $x\equiv 8\pmod 9$

13. 令 $m'=m/\gcd(c, m)$。由于 m 和 c 的公因子在求 m' 时都消去了，所以 m' 和 c 互素。由于 m 能整除 $ac-bc=(a-b)c$，所以 m' 整除 $(a-b)c$，由引理 1 知 m' 整除 $a-b$，所以 $a\equiv b\pmod{m'}$。

15. 假定 $x^2\equiv 1\pmod p$，那么 p 整除 $x^2-1=(x+1)(x-1)$。由引理 2 知 $p\mid(x+1)$ 或 $p\mid(x-1)$，所以 $x\equiv -1\pmod p$ 或 $x\equiv 1\pmod p$。

17. a) 假定 $ia\equiv ja\pmod p$，其中 $1\leqslant i<j<p$。于是 p 整除 $ja-ia=a(j-i)$。由定理 1 及 a 不能被 p 整除，p 能整除 $j-i$。但这不可能，因为 $j-i$ 是比 p 小的正整数。
   b) 由(a)知 a，$2a$，…，$(p-1)a$ 中没有两个是模 p 同余的，因此各自必定与 1 到 $p-1$ 中的不同的数同余。于是 $a\cdot 2a\cdot 3a\cdots(p-1)a\equiv 1\cdot 2\cdot 3\cdots(p-1)\pmod p$，即 $(p-1)!\cdot a^{p-1}\equiv(p-1)!\pmod p$。
   c) 根据威尔逊定理和(b)，若 p 不能整除 a，必有 $(-1)\cdot a^{p-1}\equiv -1\pmod p$。因此 $a^{p-1}\equiv 1\pmod p$。
   d) 若 $p\mid a$，那么 $p\mid a^p$，因此 $a^p\equiv a\equiv 0\pmod p$。如果 p 不能整除 a，那么由(c)知 $a^{p-1}\equiv 1\pmod p$，两边同乘以 a，给出 $a^p\equiv a\pmod p$。

19. 所有形如 $323+330k$ 的整数，其中 k 是整数。

21. 所有形如 $16+252k$ 的整数，其中 k 是整数。

23. 假定 p 是出现在 $m_1m_2\cdots m_n$ 的素数分解中的一个素数。由于这些 m_i 是两两互素的，p 只能是其中一个因子，例如是 m_j 的因子。由于 m_j 整除 $a-b$，所以 $a-b$ 的素数分解中有 p，而且 p 的次数至少是 m_j 的分解中 p 的次数。由此可知 $m_1m_2\cdots m_n$ 整除 $a-b$，所以 $a\equiv b\pmod{m_1m_2\cdots m_n}$。

25. $x\equiv 1\pmod 6$

27. a)由费马小定理知 $2^{10}\equiv 1 \pmod{11}$，因此 $2^{340}=(2^{10})^{34}\equiv 1^{34}=1 \pmod{11}$。
b)由于 $32\equiv 1 \pmod{31}$，所以 $2^{340}=(2^5)^{68}=32^{68}\equiv 1^{68}=1 \pmod{31}$。
c)由于 11 和 31 互素，$11\cdot 31=341$，由(a)，(b)和练习 19 知 $2^{340}\equiv 1 \pmod{341}$。
29. a)3，4，8　b)983
31. 首先，$2\,047=23\cdot 89$ 是合数。写 $2\,047-1=2\,046=2\cdot 1023$，所以在定义中 $s=1$ 且 $t=1\,023$。于是 $2^{2\,013}=(2^{11})^{93}=2\,048^{93}\equiv 1^{93}=1 \pmod{2\,047}$，这正是所要的。
33. 必须证明：对所有与 2 821 互素的 b，有 $b^{2\,820}\equiv 1 \pmod{2\,821}$。注意 $2\,821=7\cdot 13\cdot 31$，如果 $\gcd(b, 2\,821)=1$，则 $\gcd(b, 7)=\gcd(b, 13)=\gcd(b, 31)=1$。用费马小定理发现 $b^6\equiv 1 \pmod 7$，$b^{12}\equiv 1 \pmod{13}$，$b^{30}\equiv 1 \pmod{31}$，所以 $b^{2\,820}\equiv (b^6)^{470}\equiv 1 \pmod 7$，$b^{2\,820}\equiv (b^{12})^{235}\equiv 1 \pmod{13}$，$b^{2\,820}\equiv (b^{30})^{94}\equiv 1 \pmod{31}$。根据练习 23(或孙子剩余定理)得出 $b^{2\,820}\equiv 1 \pmod{2\,821}$，这正是所要的。
35. a)如果把这个表达式乘开，就得到 $n=1\,296m^3+396m^2+36m+1$。显然 $6m \mid n-1$，$12m \mid n-1$，$18m \mid n-1$。因此，满足练习 34 的条件，就得出 n 是卡米切尔数。
b)令 $m=51$ 就给出 $n=172\,947\,529$。
37. 0=(0，0)，1=(1，1)，2=(2，2)，3=(0，3)，4=(1，4)，5=(2，0)，6=(0，1)，7=(1，2)，8=(2，3)，9=(0，4)，10=(1，0)，11=(2，1)，12=(0，2)，13=(1，3)，14=(2，4)。
39. 我们有 $m_1=99$，$m_2=98$，$m_3=97$ 和 $m_4=95$，所以 $m=99\cdot 98\cdot 97\cdot 95=89\,403\,930$。进而 $M_1=m/m_1=903\,070$，$M_2=m/m_2=912\,285$，$M_3=m/m_3=921\,690$，$M_4=m/m_4=941\,094$。用欧几里得算法算出，对 $k=1, 2, 3, 4$ 分别有 $y_1=37$，$y_2=33$，$y_3=24$ 和 $y_4=4$ 为 M_k 模 m_k 的逆。于是解为 $65\cdot 903\,070\cdot 37+2\cdot 912\,285\cdot 33+51\cdot 921\,690\cdot 24+10\cdot 941\,094\cdot 4=3\,397\,886\,480\equiv 537\,140 \pmod{89\,403\,930}$。
41. 由练习 40 知 $\gcd(2^b-1, (2^a-1) \bmod (2^b-1))=\gcd(2^b-1, 2^{a \bmod b}-1)$。由于计算中涉及的指数是 b 和 $a \bmod b$，这与计算 $\gcd(a, b)$时涉及的量相同；用欧几里得算法计算 $\gcd(2^a-1, 2^b-1)$所使用的步骤与计算 $\gcd(a, b)$的步骤平行执行，这就证明了$\gcd(2^a-1, 2^b-1)=2^{\gcd(a,b)}-1$。
43. 假定 q 是个奇素数且 $q \mid 2^p-1$。由费马小定理知 $q \mid 2^{q-1}-1$。从练习 41 知 $\gcd(2^p-1, 2^{q-1}-1)=2^{\gcd(p,q-1)}-1$。由于 q 是 2^p-1 和 $2^{q-1}-1$ 的公因数，$\gcd(2^p-1, 2^{q-1}-1)>1$。于是 $\gcd(p, q-1)=p$，因为唯一的其他可能性即 $\gcd(p, q-1)=1$，给出 $\gcd(2^p-1, 2^{q-1}-1)=1$。因此，$p \mid q-1$，因此存在正整数 m 使 $q-1=mp$。由于 q 是奇数，m 必定是偶数，比如说 $m=2k$，于是 2^p-1 的每个素因数都是 $2kp+1$ 的形式，而且这种形式的数之积仍是这种形式。因此 2^p-1 的所有因数都是这种形式的。
45. 假定知道 $n=pq$ 和 $(p-1)(q-1)$。要计算 p 和 q，首先注意 $(p-1)(q-1)=pq-p-q+1=n-(p+q)+1$。由此可求出 $s=p+q$。由于 $q=s-p$，所以 $n=p(s-p)$。因此 $p^2-ps+n=0$。现在可以解二次方程来求 p。一旦求出 p，就可以求出 q，因为 $q=n/p$。
47. SILVER
49. $34\cdot 144+(-55)\cdot 89=1$
51.
```
procedure extended Euclidean(a, b: 正整数) x := a
y := b
oldolds := 1
olds := 0
oldoldt := 0
oldt := 1
while y≠0
begin
   q := x div y
   r := x mod y
   x := y
   y := r
   s := oldolds − q · olds
```

$t := oldoldt - q \cdot oldt$

$oldolds := olds$

$oldoldt := oldt$

$olds := s$

$oldt := t$

end{gcd(a, b)是 x，且$(oldolds)a + (oldoldt)b = x$}

53. 假定 s 是 $x^2 \equiv a(\text{mod } p)$的一个解。由于$(-s)^2 = s^2$，$-s$ 也是一个解，而且 $s \not\equiv -s(\text{mod } p)$。因为否则 $p \mid 2s$，从而 $p \mid s$，根据 s 是解的假设，$p \mid s$ 意味着 $p \mid a$，矛盾。再假定 s 和 t 是模 p 不同余的两个解，那么由于 $s^2 \equiv t^2(\text{mod } p)$，$p \mid (s^2 - t^2)$，即 $p \mid (s+t)(s-t)$。由引理 2，$p \mid (s-t)$或 $p \mid (s+t)$，所以$s \equiv t(\text{mod } p)$，或 $s \equiv -t(\text{mod } p)$。所以至多有两个解。

55. $\left(\frac{a}{p}\right)$的值只依赖于 a 是不是模 p 的二次留数，即 $x^2 \equiv a(\text{mod } p)$是否有解。由于这只与 a 模 p 的等价类有关，所以$\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$，只要 $a \equiv b(\text{mod } p)$。

57. 由练习 56 知

$$\left(\frac{a}{p}\right)\left(\frac{b}{p}\right) = a^{(p-1)/2} b^{(p-1)/2} = (ab)^{(p-1)/2} \equiv \left(\frac{ab}{p}\right)(\text{mod } p)。$$

59. $x \equiv 8, 13, 22, 27(\text{mod} 35)$

2.7 节

1. a) 3×4 b) $\begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$ c) $[2\ \ 0\ \ 4\ \ 6]$ d) 1 e) $\begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & 1 \\ 1 & 4 & 3 \\ 3 & 6 & 7 \end{bmatrix}$

3. a) $\begin{bmatrix} 1 & 11 \\ 2 & 18 \end{bmatrix}$ b) $\begin{bmatrix} 2 & -2 & -3 \\ 1 & 0 & 2 \\ 9 & -4 & 4 \end{bmatrix}$ c) $\begin{bmatrix} -4 & 15 & -4 & 1 \\ -3 & 10 & 2 & -3 \\ 0 & 2 & -8 & 6 \\ 1 & -8 & 18 & -13 \end{bmatrix}$

5. $\begin{bmatrix} 9/5 & -6/5 \\ -1/5 & 4/5 \end{bmatrix}$

7. $\mathbf{0} + \mathbf{A} = [0 + a_{ij}] = [a_{ij} + 0] = \mathbf{0} + \mathbf{A}$

9. $$\begin{aligned} \mathbf{A} + (\mathbf{B} + \mathbf{C}) &= [a_{ij} + (b_{ij} + c_{ij})] \\ &= [(a_{ij} + b_{ij}) + c_{ij}] \\ &= (\mathbf{A} + \mathbf{B}) + \mathbf{C} \end{aligned}$$

11. **A** 的行数等于 **B** 的列数，而 **A** 的列数等于 **B** 的行数。

13. $$\begin{aligned} \mathbf{A}(\mathbf{BC}) &= \left[\sum_q a_{iq}\left(\sum_r b_{qr} c_{rl}\right)\right] \\ &= \left[\sum_q \sum_r a_{iq} b_{qr} c_{rl}\right] \\ &= \left[\sum_r \sum_q a_{iq} b_{qr} c_{rl}\right] \\ &= \left[\sum_r \left(\sum_q a_{iq} b_{qr}\right) c_{rl}\right] \\ &= (\mathbf{AB})\mathbf{C} \end{aligned}$$

15. $\mathbf{A}^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$

17. a) 令 $\mathbf{A} = [a_{ij}]$，$\mathbf{B} = [b_{ij}]$，那么$(\mathbf{A} + \mathbf{B}) = [a_{ij} + b_{ij}]$。于是有$(\mathbf{A} + \mathbf{B})^t = [a_{ji} + b_{ji}] = [a_{ji}] + [b_{ji}] =$

$\mathbf{A}^t+\mathbf{B}^t$。

b) 用(a)中同样的记号，就有 $\mathbf{B}^t\mathbf{A}^t=\left[\sum_q b_{qi}a_{jq}\right]=\left[\sum_q a_{jq}b_{qi}\right]=(\mathbf{AB})^t$，这是因为第$(i,j)$项是$\mathbf{AB}$的第$(j,i)$项。

19. 结论成立是由于
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}=\begin{bmatrix} ad-bc & 0 \\ 0 & ad-bc \end{bmatrix}=(ad-bc)\mathbf{I}_2=\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

21. 由结合律，有 $\mathbf{A}^n(\mathbf{A}^{-1})^n=\mathbf{A}(\mathbf{A}\cdots(\mathbf{A}(\mathbf{A}\mathbf{A}^{-1})\mathbf{A}^{-1})\cdots\mathbf{A}^{-1})\mathbf{A}^{-1}$。由于 $\mathbf{A}\mathbf{A}^{-1}=\mathbf{I}$，从内向外乘可知 $\mathbf{A}^n(\mathbf{A}^{-1})^n=\mathbf{I}$。类似地，$(\mathbf{A}^{-1})^n\mathbf{A}^n=\mathbf{I}$。因此$(\mathbf{A}^n)^{-1}=(\mathbf{A}^{-1})^n$。

23. 需用 m_2 次乘法来计算乘积中 m_1m_3 个元素中的每一个，所以共有 $m_1m_2m_3$ 次乘法。

25. $\mathbf{A}_1((\mathbf{A}_2\mathbf{A}_3)\mathbf{A}_4)$

27. $x_1=1$，$x_2=-1$，$x_3=-2$

29. a) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$

31. a) $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$ c) $\begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

33. a) $\mathbf{A}\vee\mathbf{B}=[a_{ij}\vee b_{ij}]=[b_{ij}\vee a_{ij}]=\mathbf{B}\vee\mathbf{A}$

b) $\mathbf{A}\wedge\mathbf{B}=[a_{ij}\wedge b_{ij}]=[b_{ij}\wedge a_{ij}]=\mathbf{B}\wedge\mathbf{A}$

35. a)
$$\begin{aligned}\mathbf{A}\vee(\mathbf{B}\wedge\mathbf{C})&=[a_{ij}]\vee[b_{ij}\wedge c_{ij}]\\&=[a_{ij}\vee(b_{ij}\wedge c_{ij})]\\&=[(a_{ij}\vee b_{ij})\wedge(a_{ij}\vee c_{ij})]\\&=[a_{ij}\vee b_{ij}]\wedge[a_{ij}\vee c_{ij}]\\&=(\mathbf{A}\vee\mathbf{B})\wedge(\mathbf{A}\vee\mathbf{C})\end{aligned}$$

b)
$$\begin{aligned}\mathbf{A}\wedge(\mathbf{B}\vee\mathbf{C})&=[a_{ij}]\wedge[b_{ij}\vee c_{ij}]\\&=[a_{ij}\wedge(b_{ij}\vee c_{ij})]\\&=[(a_{ij}\wedge b_{ij})\vee(c_{ij}\wedge c_{ij})]\\&=[a_{ij}\wedge b_{ij}]\vee[a_{ij}\wedge c_{ij}]\\&=(\mathbf{A}\wedge\mathbf{B})\vee(\mathbf{A}\wedge\mathbf{C})\end{aligned}$$

37.
$$\begin{aligned}\mathbf{A}\odot(\mathbf{B}\odot\mathbf{C})&=\left[\bigvee_q a_{iq}\wedge\left(\bigvee_r (b_{qr}\wedge c_{rl})\right)\right]\\&=\left[\bigvee_q\bigvee_r (a_{iq}\wedge b_{qr}\wedge c_{rl})\right]\\&=\left[\bigvee_r\bigvee_q (a_{iq}\wedge b_{qr}\wedge c_{rl})\right]\\&=\left[\bigvee_r\left(\bigvee_q (a_{iq}\wedge b_{qr})\right)\wedge c_{rl}\right]\\&=(\mathbf{A}\odot\mathbf{B})\odot\mathbf{C}\end{aligned}$$

补充练习

1. a) **procedure** *last max*$(a_1, \cdots, a_n$: 整数)

$max := a_1$

$last := 1$

```
    i := 2
    while i ≤ n
    begin
        if a_i ≥ max then
        begin
            max := a_i
            last := i
        end
        i := i + 1
    end {last 是表中最大整数的最末次出现位置}
```

b) $2n-1=O(n)$ 次比较。

3\. a)

```
procedure pair zeros(b_1 b_2 … b_n: 位串, n ≥ 2)
    x := b_1
    y := b_2
    k := 2
    while (k < n 且(x ≠ 0 或 y ≠ 0))
    begin
        k := k + 1
        x := y
        y := b_k
    end
    if(x = 0 且 y = 0) then print“有”
    else print “无”
```

b) $O(n)$ 次比较。

5\. a) 和 b)

```
procedure smallest and largest(a_1, a_2, …, a_n: 整数)
min := a_1
max := a_1
for i := 2 to n
begin
    if a_i < min then min := a_i
    if a_i > max then max := a_i
end {min 是输入中最小整数，max 是最大整数}
```

c) $2n-2$

7\. 在每个元素的任意一次比较之前都有这样的可能性，即这个元素可能是最大的，也可能是最小的。这意味着有 $2n$ 种不同的可能性，需要通过元素比较来消除其中的 $2n-2$ 种可能，因为需要找出唯一的最大元和唯一的最小元。根据所比较的两个元素是否已经参与过比较，把两个元素的比较分成“首次”和“非首次”。首次比较排除了较大元素是最小元和较小元素是最大元的可能性。因此每个首次比较消除了两种可能性，显然不能消除更多种可能性了。非首次比较一定是在两个都有可能成为最大元或两个都有可能成为最小元的元素之间进行的，其中至少有一个元素一定不可能成为相反方向的最×元。比如可能比较 x 和 y，其中只知道 x 不可能是最小元。如果在这种情形下发现 $x>y$，则只消除了一种可能性，即 y 不是最大元。因此在最坏情形下，一次非首次比较只能消除一种可能性。(其他非首次比较的情形是类似的。)于是至多有 $\lfloor n/2 \rfloor$ 次没有比较过的元素的比较，每次消除两种可能，总共消除 $2\lfloor n/2 \rfloor$ 种可能。因此为了求出答案，还需要 $2n-2-2\lfloor n/2 \rfloor$ 次比较，因为已经论证过，这些比较在最坏情形下只能消除一

种可能，而总共有 $2n-2$ 种可能。但是 $2n-2-2\lfloor n/2\rfloor+\lfloor n/2\rfloor=2n-2-\lfloor n/2\rfloor=2n-2+\lceil -n/2\rceil=\lceil 2n-n/2\rceil-2=\lceil 3n/2\rceil-2$，这正是需要的结果。

9. 第一遍结束时：3，1，4，5，2，6；第二遍结束时：1，3，2，4，5，6；第三遍结束时：1，2，3，4，5，6；第四遍发现无须交换，算法结束。

11. 至多可能对表做 n 遍扫描，每遍有 n 次比较。因此总共有 $O(n^2)$ 次比较。

13. 由于对所有 $n>0$ 有 $\log n<n$，所以 $(n\log n+n^2)^3\leqslant(n^2+n^2)^3\leqslant(2n^2)^3=8n^6$，让见证 $C=8$ 和 $k=0$ 就证明 $(n\log n+n^2)^3=O(n^6)$。

15. $O(x^2 2^x)$

17. 注意 $\frac{n!}{2^n}=\frac{n}{2}\cdot\frac{n-1}{2}\cdots\frac{3}{2}\cdot\frac{2}{2}\cdot\frac{1}{2}>\frac{n}{2}\cdot 1\cdot 1\cdots 1\cdot\frac{1}{2}=\frac{n}{4}$。

19. 5，22，-12，-29

21. 由于 $ac\equiv bc\pmod m$，所以存在整数 k 使 $ac=bc+km$。因此 $a-b=km/c$。由于 $a-b$ 是整数，所以 $c\mid km$。令 $d=\gcd(m,c)$，写成 $c=de$。由于 e 没有能整除 m/d 的因子，所以 $d\mid m$ 且 $e\mid k$。因此 $a-b=(k/e)(m/d)$，其中 $k/e\in\mathbf{Z}$ 且 $m/d\in\mathbf{Z}$。于是 $a\equiv b\pmod{m/d}$。

23. 1

25. 1

27. $(a_n a_{n-1}\cdots a_1 a_0)_{10}=\sum_{k=0}^{n}10^k\cdot a_k\equiv\sum_{k=0}^{n}a_k\pmod 9$，因为对每个非负整数 k 均有 $10^k\equiv 1\pmod 9$。

29. a)不互素　　b)互素　　c)互素　　d)互素

31. a)解密函数是 $g(q)=\bar{a}(q-b)\bmod 26$，其中 $\bar{a}$ 是 a 模 26 的一个逆。
b)PLEASE SEND MONEY

33. $x\equiv 28\pmod{30}$

35. 对 $n\geqslant 0$，$\mathbf{A}^{4n}=\begin{bmatrix}1&0\\0&1\end{bmatrix}$，$\mathbf{A}^{4n+1}=\begin{bmatrix}0&1\\-1&0\end{bmatrix}$，$\mathbf{A}^{4n+2}=\begin{bmatrix}-1&0\\0&-1\end{bmatrix}$，$\mathbf{A}^{4n+3}=\begin{bmatrix}0&-1\\1&0\end{bmatrix}$。

37. 假定 $\mathbf{A}=\begin{bmatrix}a&b\\c&d\end{bmatrix}$。令 $\mathbf{B}=\begin{bmatrix}0&1\\0&0\end{bmatrix}$，由于 $\mathbf{AB}=\mathbf{BA}$，所以 $c=0$ 而 $a=d$。令 $\mathbf{B}=\begin{bmatrix}0&0\\1&0\end{bmatrix}$，由于 $\mathbf{AB}=\mathbf{BA}$，所以 $b=0$。因此 $\mathbf{A}=\begin{bmatrix}a&0\\0&a\end{bmatrix}=a\mathbf{I}$。

39. **procedure** *triangular matrix multiplication*($\mathbf{A}$，$\mathbf{B}$：$n\times n$ 上三角矩阵，$\mathbf{A}=[a_{ij}]$，$\mathbf{B}=[b_{ij}]$)

```
for i := 1 to n
begin
    for j := i to n
    begin
        c_ij := 0
        for k := i to j
            c_ij := c_ij + a_ik b_kj
    end
end
```

41. $(\mathbf{AB})(\mathbf{B}^{-1}\mathbf{A}^{-1})=\mathbf{A}(\mathbf{BB}^{-1})\mathbf{A}^{-1}=\mathbf{AIA}^{-1}=\mathbf{AA}^{-1}=\mathbf{I}$。同理，$(\mathbf{B}^{-1}\mathbf{A}^{-1})(\mathbf{AB})=\mathbf{I}$。所以 $(\mathbf{AB})^{-1}=\mathbf{B}^{-1}\mathbf{A}^{-1}$。

43. a)令 $\mathbf{A}\odot\mathbf{0}=[b_{ij}]$，则 $b_{ij}=(a_{i1}\wedge 0)\vee\cdots\vee(a_{ip}\wedge 0)=0$。因此 $\mathbf{A}\odot\mathbf{0}=\mathbf{0}$。同理，$\mathbf{0}\odot\mathbf{A}=\mathbf{0}$。
b)$\mathbf{A}\vee\mathbf{0}=[a_{ij}\vee 0]=[a_{ij}]=\mathbf{A}$，因此 $\mathbf{A}\vee\mathbf{0}=\mathbf{A}$。同理，$\mathbf{0}\vee\mathbf{A}=\mathbf{A}$。
c)$\mathbf{A}\wedge\mathbf{0}=[a_{ij}\wedge 0]=[0]=\mathbf{0}$，因此 $\mathbf{A}\wedge\mathbf{0}=\mathbf{0}$。同理，$\mathbf{0}\wedge\mathbf{A}=\mathbf{0}$。

45. 假设有人选择了一个小于 2^n 的正整数，我们打算猜出这个数。让这个人把这个数写成二进制，必要时

在开头加 0 使之有 n 位长。于是问："第一位是 1 吗?"，"第二位是 1 吗?"，"第三位是 1 吗?"，依次类推。在知道了这 n 个问题的答案后，就知道了这个数，因为知道了这个数的二进制展开。

第 3 章

3.1 节

1. 由于相邻二整数之积能被 2 整除，故所论之积能被 2 整除。由于相邻三整数之积能被 3 整除，故所论之积能被 3 整除。因此所论之积的素因子分解中含有 2 和 3，所以能被 $2\cdot 3$ 整除。
3. $n^4-1=(2k+1)^4-1=16k^4+32k^3+24k^2+8k=8k(k+1)(2k^2+2k+1)$。$k$ 或 $k+1$ 之一是偶数，所以 16 能整除 n^4-1。
5. 由于 $5^4=625$，所以 x 和 y 都应小于 5。于是 $x^4+y^4\leqslant 4^4+4^4=512<625$。
7. 如果 $3x^2-1=8y$，则 $3x^2\equiv 1(\bmod 8y)$。如果 x 是奇数，则由练习 2，$x^2\equiv 1(\bmod 8)$，所以 $3x^2\equiv 3(\bmod 8)$。如果 x 是偶数，则 $x^2\equiv 0$ 或 $4(\bmod 8)$，所以 $3x^2\equiv 0$ 或 $4(\bmod 8)$。这个矛盾说明题目中的方程无整数解。
9. 如果 $x^4-3=16y^4$，则 $x^4\equiv 3(\bmod 16)$。根据练习 3，偶数的 4 次幂模 16 同余 0。因此 x^4 不可能模 16 同余 3，所以该方程无解。
11. 不同正实数 a 和 b 的调和均值总是小于几何均值。为了证明 $2ab/(a+b)<\sqrt{ab}$，两边同乘以 $(a+b)/(2\sqrt{ab})$，得到等价的不等式 $\sqrt{ab}<(a+b)/2$，例 1 证明过这个不等式。
13. 写在黑板上数字之和的奇偶性是不会改变的，因为 $j+k$ 与 $|j-k|$ 同奇偶(在每一步这个和减少 $j+k$ 而增加 $|j-k|$)。因此过程结束时的整数必定与 $1+2+\cdots+(2n)=n(2n+1)$ 同奇偶，由于 n 是奇数，$n(2n+1)$ 是奇数。
15. $n=1\,601$ 是一个反例。
17. 假设 $3^{1/3}=a/b$，其中 a，$b\in\mathbf{Z}$，$b\neq 0$，而且 $\gcd(a,b)=1$。于是 $3=a^3/b^3$，所以 $3b^3=a^3$。因此 $3\mid a^3$，仅当 $3\mid a$ 时才可以发生这样的事。设 $a=3m$，则 $3b^3=27m^3$，或 $b^3=9m^3$。因此 $3\mid b^3$，这说明 $3\mid b$。这与假设 $\gcd(a,b)=1$ 矛盾。
19. 根据高中代数的有理根判别法，任何满足这个整系数多项式方程的有理数都形如 p/q，其中 p 是常数项(1)的因子，q 是首项系数(1)的因子。所以仅有的可能解是 ± 1；由于 ± 1 都不是根，所以无有理根。
21. $a^2\equiv b^2(\bmod p)$ 当且仅当 $p\mid(a^2-b^2)=(a+b)(a-b)$。根据素因子分解的唯一性，这等价于 $p\mid(a-b)$ 或 $p\mid(a+b)$，这与 $a\equiv b(\bmod p)$ 或 $a\equiv -b(\bmod p)$ 是一样的。
23. 两边都是小于 m 的非负整数，都模 m 同余 ab，所以两边相等。
25. 3，5，7 即所求的素数。
27. 假设有前 $n+1$ 个素数 p_1，p_2，…，p_{n+1}。则 $p_1p_2\cdots p_{n+1}$ 被超过 n 个素数整除。
29. 若不存在，则假设 q_1，q_2，…，q_n 是所有形如 $6k+5$ 的素数。设 $Q=6q_1q_2\cdots q_n-1$。注意 Q 形如 $6k+5$，其中 $k=q_1q_2\cdots q_n-1$。设 Q 的素因子分解是 $Q=p_1p_2\cdots p_t$。各个 p_i 都不是 2，3 或 q_j，因为 Q 除以 2 余 1，除以 3 余 2，除以 q_j 余 q_j-1。除 3 以外的奇素数都形如 $6k+1$ 或 $6k+5$，形如 $6k+5$ 的素数之积仍然形如 $6k+5$。因此至少有一个 p_i 必定形如 $6k+5$，矛盾。
31. 通过考虑下列 3 种情形来证明：对于某个整数 k，$n=3k$，$n=3k+1$，$n=3k+2$。若 n 是 3 的倍数，则左边是 $\left\lfloor\frac{x}{3}\right\rfloor+\left\lfloor\frac{x+n}{3}\right\rfloor+\left\lfloor\frac{x+2n}{3}\right\rfloor=\left\lfloor\frac{x}{3}\right\rfloor+\left\lfloor\frac{x}{3}\right\rfloor+\frac{n}{3}+\left\lfloor\frac{x}{3}\right\rfloor+\frac{2n}{3}=3\left\lfloor\frac{x}{3}\right\rfloor+n$，等式右边也是 $3\left\lfloor\frac{x}{3}\right\rfloor+n$。若 $n=3k+1$，则左边是 $\left\lfloor\frac{x}{3}\right\rfloor+\left\lfloor\frac{x+3k+1}{3}\right\rfloor+\left\lfloor\frac{x+6k+2}{3}\right\rfloor=\left\lfloor\frac{x}{3}\right\rfloor+\left\lfloor\frac{x+1}{3}\right\rfloor+k+\left\lfloor\frac{x+2}{3}\right\rfloor+2k=\left\lfloor\frac{x}{3}\right\rfloor+\left\lfloor\frac{x+1}{3}\right\rfloor+\left\lfloor\frac{x+2}{3}\right\rfloor+n-1$，右边是 $\lfloor x\rfloor+n-1$。

根据 1.8 节练习 68，这两边是相等的。第三种情形是类似的。

33. a)不失一般性，假设 x 序列已经排成非降顺序，因为可以重新标记下标。y 序列只有有穷多种可能的顺序，所以如果能够证明，每当发现逆序的 y_i 和 y_j(即 $i<j$ 且 $y_i>y_j$)时，都可以通过交换它们来让和增大(或至少保持不变)，则就证明了当 y 序列排成非降顺序时，和最大。其实，如果做这样的交换，和就增加 $x_iy_j+x_jy_i$ 而减少 $x_iy_i+x_jy_j$。净效果是增加了 $x_iy_j+x_jy_i-x_iy_i-x_jy_j=(x_j-x_i)(y_i-y_j)$，根据序列顺序的假设，这个值是非负的。

    b)与 a)类似。

35. 令 $x=n+(r/m)+\varepsilon$，其中 n 是整数，r 是小于 m 的非负整数，ε 是实数，$0\leqslant\varepsilon\leqslant 1/m$。左边是 $\lfloor nm+r+m\varepsilon\rfloor=nm+r$。在右边，从 $\lfloor x\rfloor$ 到 $\lfloor x+(m+r-1)/m\rfloor$ 的项都恰好是 n，从 $\lfloor x+(m+r)/m\rfloor$ 往右的项都恰好是 $n+1$。因此右边也是 $(m-r)n+r(n+1)=nm+r$。

37. 回忆一下，非常值多项式只能有穷多次取相同的值。因此 f 只能有穷多次取 0 和 ±1，所以如果不存在整数 y 使得 $f(y)$ 是合数，则必定存在某个 x_0 使得 $\pm f(x_0)$ 是素数，比如说 p。考虑 $f(x_0+kp)$，当在该多项式中用 x_0+kp 代入 x 并展开时，除了构成 $f(x_0)$ 的那些项之外，每项都含有因子 p。因此对某个整数 m，有 $f(x_0+kp)=f(x_0)+mp=(m+1)p$。当 k 变化时，这个值只能有穷多次取 0，p 或 $-p$；因此对某个 k 值，它一定是合数。

39. 假设每个大于 2 的偶数都是两个素数之和，设 n 是大于 5 的整数。若 n 是奇数，则写 $n=3+(n-3)$ 并且分解 $n-3=p+q$ 为两个素数之和；若 n 是偶数，则写 $n=2+(n-2)$ 并且分解 $n-2=p+q$ 为两个素数之和。反之，假设每个大于 5 的整数都是三个素数之和，设 n 是大于 2 的偶数。把 $n+2$ 写成三个素数之和，其中一个必定是 2，所以 $n+2=2+p+q$，所以 $n=p+q$。

41. a)$6\to3\to10\to5\to16\to8\to4\to2\to1$

    b)$7\to22\to11\to34\to17\to52\to26\to13\to40\to20\to10\to5\to16\to8\to4\to2\to1$

    c)$17\to52\to26\to13\to40\to20\to10\to5\to16\to8\to4\to2\to1$

    d)$21\to64\to32\to16\to8\to4\to2\to1$

43. $1+2+3=6$；$1+2+4+7+14=28$；$1+2+4+8+16+31+62+124+248=496$。

45. 真因子之和是$(1+2+4+8+\cdots+2^{p-1})+(2^p-1)(1+2+4+8+\cdots+2^{p-2})=2^p-1+(2^p-1)(2^{p-1}-1)=(2^p-1)(1+2^{p-1}-1)=(2^p-1)2^{p-1}$。

47. $(3,5,0)\to(3,2,3)\to(6,2,0)\to(6,0,2)\to(1,5,2)\to(1,4,3)$

49. 假设存在程序 S 能判定带有给定输入的程序是否会打印数字 1。于是就有一个解停机问题的算法：给定程序 P 及其输入 I，构造程序 P'，P' 与 P 基本一样，但是 P' 不打印任何东西(即使 P 打印一些东西)，当且仅当 P' 要停机时，P' 才打印 1 并且停机。于是 P 在一个输入上停机当且仅当 P' 在同样的输入上打印 1。把 P' 和 I 输入到 S，S 将判定 P 是否在输入 I 上停机。由于已知停机问题是不可解的，所以这是矛盾。因此这样的程序 S 不存在。

51. 并行运行这两个程序，等待其中之一停机。

3.2 节

1. a)3　　b)-1　　c)787　　d)2 639

3. a)$a_0=2$，$a_1=3$，$a_2=5$，$a_3=9$　　b)$a_0=1$，$a_1=4$，$a_2=27$，$a_3=256$

    c)$a_0=0$，$a_1=0$，$a_2=1$，$a_3=1$　　d)$a_0=0$，$a_1=1$，$a_2=2$，$a_3=3$

5. a)2，5，8，11，14，17，20，23，26，29

    b)1，1，1，2，2，2，3，3，3，4

    c)1，1，3，3，5，5，7，7，9，9

    d)-1，-2，-2，8，88，656，4 912，40 064，362 368，3 627 776

    e)3，6，12，24，48，96，192，384，768，1 536

f)1，1，2，3，5，8，13，21，34，55

g)1，2，2，3，3，3，3，4，4，4

h)3，3，5，4，4，3，5，5，4，3

7. 每一项可以是前一项的两倍；第 n 项可以从前一项加上 $n-1$ 得到；这些项可以是并非 3 的倍数的正整数；还有无穷多种其他的可能性。

9. a)1 个 1 和 1 个 0，再跟上 2 个 1 和 2 个 0，再跟上 3 个 1 和 3 个 0，等等。

b)正整数按增序排列，而且偶数出现两次。

c)奇数位的项是 2 的一次次幂；偶数位的项全是 0。

d)$a_n=3\cdot 2^{n-1}$　　e)$a_n=15-7(n-1)=22-7n$

f)$a_n=(n^2+n+4)/2$　　g)$a_n=2n^3$　　h)$a_n=n!+1$

11. 由于 a_n 表示第 n 个不是完全平方的正整数，在 1，2，…，a_n 中，a_1，a_2，…，a_n 是那些不为完全平方的数，1^2，2^2，…，k^2 是为完全平方的数，其中 k 是使 $k^2<n+k<(k+1)^2$ 的整数。于是 $a_n<n+k$，$k^2<a_n<(k+1)^2$。为求 k，首先注意 $k^2<n+k<(k+1)^2$，所以，$k^2+1\leqslant n+k\leqslant(k+1)^2-1$。因而 $\left(k-\frac{1}{2}\right)^2+\frac{3}{4}=k^2-k+1\leqslant n\leqslant k^2+k=\left(k+\frac{1}{2}\right)^2-\frac{1}{4}$。由此可知 $k-\frac{1}{2}<\sqrt{n}<k+\frac{1}{2}$，所以 $k=\{\sqrt{n}\}$，$a_n=n+k=n+\{\sqrt{n}\}$。

13. a)20　b)11　c)30　d)511

15. a)1 533　b)510　c)4 923　d)9 842

17. a)21　b)78　c)18　d)18

19. $\sum_{j=1}^{n}(a_j-a_{j-1})=a_n-a_0$

21. a)n^2　b)$n(n+1)/2$

23. 15 150

25. $\frac{n(n+1)(2n+1)}{3}+\frac{n(n+1)}{2}+(n+1)(m-(n+1)^2+1)$其中 $n=\lfloor\sqrt{m}\rfloor-1$

27. a)0　b)1 680　c)1　d)1 024

29. 34

31. a)可数，-1，-2，-3，-4，…　b)可数，0，2，-2，4，-4，…

c)不可数。　d)可数，0，7，-7，14，-14，…

33. 假定 $A-B$ 可数，那么由于 $A=(A-B)\cup B$，A 中元素可以用交替列出 $A-B$ 中元素和 B 中元素的方式排成序列。这和 A 不可数矛盾。

35. 假定 B 可数，那么 B 中元素可以列成 b_1，b_2，b_3，…。由于 A 是 B 的子集，取该序列的子序列$\{b_n\}$使它包含属于 A 的项，就得到 A 中元素的一个序列。但 A 是不可数的，不可能排成序列。

37. 假定 $A_1,A_2,A_3,\cdots$ 全是可数集合。由于 A_i 是可数的，可以把它的元素排成序列 $a_{i1},a_{i2},a_{i3},\cdots$。集合 $\bigcup_{i=1}^{n}A_i$ 的元素 a_{ij} 可以按下列方式排成序列：先排 $i+j=2$ 的元素，再排 $i+j=3$ 的元素，然后是 $i+j=4$ 的元素，等等。

39. 长度为 m 的位串只有有限多个，即 2^m 个。所有位串的集合是长度为 m 的位串的集合的并集，$m=0$，1，2，…。由于可数多个可数集合的并集是可数的，所以只有可数多个位串。

41. 对任何有限的字母表，只有有限多个长度为 n 的位串，其中 n 为正整数。从练习 37 可知，有限字母表上长度不超过给定正整数的位串只有有限多个。由于用特定程序语言写的计算机程序的集合是某个有限字母表上的所有位串之集合的子集，而根据练习 34，所有位串的集合是可数的，所以作为子集的程序集合是可数的。

43. 练习 41 证明了只有可数多个计算机程序。因此只有可数多个可计算的函数。据练习 42 所示，函数个

数是不可数的，所以并非所有函数都是可计算的。

45. 给定正整数 x，证明恰好存在一个正有理数 m/n（约除公因子），使得 $K(m/n)=x$。从 x 的素因子分解读出 m/n，使得 $K(m/n)=x$。出现在偶次幂中的素数都出现在 m 的素因子分解中，但指数减半。出现在奇次幂中的素数都出现在 n 的素因子分解中，但指数加 1 减半。

3.3 节

1. $n(n+1)$

3. 设 $P(n)$ 是“$\sum_{j=0}^{n} 3\cdot 5^j=3(5^{n+1}-1)/4$”。

基础步骤：$P(0)$ 为真，因为 $\sum_{j=0}^{0} 3\cdot 5^j=3=3(5^1-1)/4$。

归纳步骤：假设 $\sum_{j=0}^{k} 3\cdot 5^j=3(5^{k+1}-1)/4$。则 $\sum_{j=0}^{k+1} 3\cdot 5^j=\left(\sum_{j=0}^{k} 3\cdot 5^j\right)+3\cdot 5^{k+1}=3(5^{k+1}-1)/4+3\cdot 5^{k+1}=3(5^{k+1}+4\cdot 5^{k+1}-1)/4=3(5^{k+2}-1)/4$。

5. 通过检查小的 n 值，做出猜想：$P(n)$ 为真，其中 $P(n)$ 是命题“$\sum_{j=1}^{n} 1/2^j=(2^n-1)/2^n$”。

基础步骤：$P(1)$ 为真，因为 $\frac{1}{2}=(2^1-1)/2^1$。

归纳步骤：假设 $\sum_{j=1}^{k} 1/2^j=(2^k-1)/2^k$，则

$$\sum_{j=1}^{k+1}\frac{1}{2^j}=\left(\sum_{j=1}^{k}\frac{1}{2^j}\right)+\frac{1}{2^{k+1}}=\frac{2^k-1}{2^k}+\frac{1}{2^{k+1}}=\frac{2^{k+1}-2+1}{2^{k+1}}=\frac{2^{k+1}-1}{2^{k+1}}$$

7. 设 $P(n)$ 是“$\sum_{j=1}^{n} j^2=n(n+1)(2n+1)/6$”。

基础步骤：$P(1)$ 为真，因为 $\sum_{1}^{1} j^2=1=1(1+1)(2\cdot 1+1)/6$。

归纳步骤：则 $\sum_{j=1}^{k+1} j^2=\left(\sum_{j=1}^{k} j^2\right)+(k+1)^2=k(k+1)(2k+1)/6+(k+1)^2=(k+1)[2k^2+k+6k+6]/6=(k+1)(k+2)\cdot(2k+3)/6=(k+1)((k+1)+1)(2(k+1)+1)/6$。

9. 设 $P(n)$ 是“$1^2+3^2+\cdots+(2n+1)^2=(n+1)(2n+1)(2n+3)/3$”。

基础步骤：$P(0)$ 为真，因为 $1^2=1=(0+1)(2\cdot 0+1)(2\cdot 0+3)/3$。

归纳步骤：假设 $P(k)$ 为真。则 $1^2+3^2+\cdots+(2k+1)^2+(2(k+1)+1)^2=(k+1)(2k+1)(2k+3)/3+(2k+3)^2=(2k+3)[(k+1)(2k+1)/3+(2k+3)]=(2k+3)(2k^2+9k+10)/3=(2k+3)(2k+5)(k+2)/3=((k+1)+1)(2(k+1)+1)(2(k+1)+3)/3$。

11. 设 $P(n)$ 是“$1+nh\leqslant(1+h)^n$，$h>-1$”。

基础步骤：$P(0)$ 为真，因为 $1+0\cdot h=1\leqslant 1=(1+h)^0$。

归纳步骤：假设 $1+kh\leqslant(1+h)^k$。则因为 $(1+h)>0$，所以 $(1+h)^{k+1}=(1+h)(1+h)^k\geqslant(1+h)(1+kh)=1+(k+1)h+kh^2\geqslant 1+(k+1)h$。

13. 设 $P(n)$ 是“$2^n>n^2$”。

基础步骤：$P(5)$ 为真，因为 $2^5=32>25=5^2$。

归纳步骤：假设 $P(k)$ 为真，即 $2^k>k^2$。则 $2^{k+1}=2\cdot 2^k>k^2+k^2>k^2+4k\geqslant k^2+2k+1=(k+1)^2$，因为 $k>4$。

15. 设 $P(n)$ 是“$1\cdot 2+2\cdot 3+\cdots+n(n+1)=n(n+1)(n+2)/3$”。

基础步骤：$P(1)$ 为真，因为 $1\cdot 2=2=1(1+1)(1+2)/3$。

归纳步骤：假设 $P(k)$ 为真。则 $1\cdot 2+2\cdot 3+\cdots+k(k+1)+(k+1)(k+2)=[k(k+1)(k+2)/3]+$

$(k+1)(k+2)=(k+1)(k+2)[(k/3)+1]=(k+1)(k+2)(k+3)/3$。

17. 设 $P(n)$ 是"$1^2-2^2+3^2-\cdots+(-1)^{n-1}n^2=(-1)^{n-1}n(n+1)/2$"。

    基础步骤：$P(1)$ 为真，因为 $1^2=1=(-1)^0 1^2$。

    归纳步骤：假设 $P(k)$ 为真。则 $1^2-2^2+3^2-\cdots+(-1)^{k-1}k^2+(-1)^k(k+1)^2=(-1)^{k-1}k(k+1)/2+(-1)^k(k+1)^2=(-1)^k(k+1)[-k/2+(k+1)]=(-1)^k(k+1)[(k/2)+1]=(-1)^k(k+1)(k+2)/2$。

19. 设 $P(n)$ 是"n 分邮资可以用 3 分和 5 分邮票来组成"。

    基础步骤：$P(8)$ 为真，因为 8 分邮资可以用一个 3 分和一个 5 分邮票来组成。

    归纳步骤：假设 $P(k)$ 为真，即可以组成 k 分邮资。将说明如何组成 $k+1$ 分邮资。根据归纳假设可以组成 k 分邮资。若这里面包含一个 5 分邮票，则用两个 3 分邮票替换这个 5 分邮票来得到 $k+1$ 分邮资。否则，只用过 3 分邮票并且 $k\geqslant 9$。删除三个 3 分邮票，用两个 5 分邮票替换它们来得到 $k+1$ 分邮资。

21. 设 $P(n)$ 是"n^5-n 被 5 整除"。

    基础步骤：$P(0)$ 为真，因为 $0^5-0=0$ 被 5 整除。

    归纳步骤：假设 $P(k)$ 为真，即 k^5-5 被 5 整除，则 $(k+1)^5-(k+1)=(k^5+5k^4+10k^3+10k^2+5k+1)-(k+1)=(k^5-k)+5(k^4+2k^3+2k^2+k)$ 也被 5 整除，因为在这个和里的两项都被 5 整除。

23. 设 $P(n)$ 是命题：$(2n-1)^2-1$ 被 8 整除。基础情形 $P(1)$ 为真，因为 $8\mid 0$。现在假设 $P(k)$ 为真。因为 $((2(k+1)-1)^2-1)=((2k-1)^2-1)+8k$，所以 $P(k+1)$ 为真，因为右边两项都被 8 整除。这说明对所有正整数来说 $P(n)$ 都为真，所以每当 m 是奇正整数时，就有 m^2-1 被 8 整除。

25. 设 $P(n)$ 是命题：带 n 个元素的集合有 $n(n-1)/2$ 个双元素子集合。基础情形 $P(2)$ 为真，因为带 2 个元素的集合有一个双元素子集合，即它自身，而且 $2(2-1)/2=1$。现在假设 $P(k)$ 为真。设 S 是带 $k+1$ 个元素的集合。选择 S 里的一个元素 a 并且设 $T=S-\{a\}$。S 的双元素子集合要么包含 a 要么不包含 a。不包含 a 的那些子集合都是 T 的 2 元素子集合；根据归纳假设，有 $k(k-1)/2$ 个这样的子集合。包含 a 的 S 的双元素子集合有 k 个，因为这样的子集合包含 a 和 T 里的 k 个元素之一。因此有 $k(k-1)/2+k=k(k+1)/2$ 个 S 的双元素子集合。这样就完成了归纳证明。

27. 设 $P(n)$ 是命题：$1^4+2^4+3^4+\cdots+n^4=n(n+1)(2n+1)(3n^2+3n-1)/30$。$P(1)$ 为真，因为 $1\cdot 2\cdot 3\cdot 5/30=1$。假设 $P(k)$ 为真。则 $(1^4+2^4+3^4+\cdots+k^4)+(k+1)^4=k(k+1)(2k+1)(3k^2+3k-1)/30+(k+1)^4=((k+1)/30)(k(2k+1)(3k^2+3k-1)+30(k+1)^3)=((k+1)/30)(6k^4+39k^3+91k^2+89k+30)=((k+1)/30)(k+2)(2k+3)(3(k+1)^2+3(k+1)-1)$。这说明 $P(k+1)$ 为真。

29. 通过检查发现，对 $n=0, 1, 2, 3$ 来说，不等式 $2n+3\leqslant 2^n$ 不成立。设 $P(n)$ 是命题：对正整数 n 来说，这个不等式成立。基础情形 $P(4)$ 为真，因为 $2\cdot 4+3=11\leqslant 16=2^4$。

    对于归纳步骤，假设 $P(k)$ 为真。则根据归纳假设，$2(k+1)+3=(2k+3)+2<2^k+2$。但是因为 $k\geqslant 1$，所以 $2^k+2\leqslant 2^k+2^k=2^{k+1}$。这说明 $P(k+1)$ 为真。

31. a)用 5 分和 6 分邮票可以组成的邮资是 5 分、6 分、10 分、11 分、12 分、15 分、16 分、17 分、18 分以及 20 分以上的所有邮资。

    b)将证明 20 分以上的所有邮资都可以用 5 分和 6 分邮票来组成。设 $P(n)$ 是命题：可以组成 n 分邮资。$P(20)$ 为真，因为 20 分邮资可以用四个 5 分邮票来组成。现在假设 $P(k)$ 为真。若使用过一个 5 分邮票来组成 n 分邮资，则用一个 6 分邮票替换它来组成 $k+1$ 分邮资。否则，若只用过 6 分邮票，则因为 $k\geqslant 20$，所以至少用过四个 6 分邮票。用五个 5 分邮票替换四个 6 分邮票来得到 $k+1$ 分邮资。因为 $P(k+1)$ 为真。这样就完成了数学归纳法证明。

    c)设 $P(n)$ 与在(b)里一样。基础情形是 $P(20)$、$P(21)$、$P(22)$、$P(23)$ 和 $P(24)$。这些情形都为真，因为 20 分、21 分、22 分、23 分和 24 分邮资分别可以用四个 5 分邮票、三个 5 分邮票和一个 6 分邮票、两个 5 分邮票和两个 6 分邮票、一个 5 分邮票和三个 6 分邮票以及四个 6 分邮票来组成。现在假设对 $20\leqslant j\leqslant k$ 来说 $P(j)$ 为真，其中 $k\geqslant 24$。因为 $k+1\geqslant 25$，所以 $k-4\geqslant 20$，所以根据归纳假设，可以组成 $k-4$ 分邮资。添加一个 5 分邮票来得到 $k+1$ 分邮资，这说明 $P(k+1)$ 为真。这样就完成了用数学归纳法第二原理的证明。

33. 所有大于或等于 40 美元的 10 美元的倍数，以及 20 美元，都可以组成。设 $P(n)$是命题：可以组成 $10n$ 美元。$P(4)$为真，因为用两个 20 美元可以组成 40 美元。现在假设对 $k\geqslant 4$ 来说 $P(k)$为真。若用过一个 50 美元钞票来组成 $10k$ 美元，则用三个 20 美元钞票替换它来得到 $10(k+1)$美元。否则，至少用过两个 20 美元钞票，因为 $10k$ 至少为 40 美元。用一个 50 美元钞票替换两个 20 美元钞票来得到 $10(k+1)$ 美元。这说明 $P(k+1)$为真。

35. *基础步骤*：有 4 种基本情形。如果 $n=1=4\cdot 0+1$，则显然选手二获胜。如果有二、三、四根火柴($n=4\cdot 0+2$，$n=4\cdot 0+3$，$n=4\cdot 1$)，则通过只留下一根火柴，选手一获胜。

归纳步骤：假设强归纳假设，即在有不超过 k 根火柴的游戏中，若 $k\equiv 0, 2, 3 \pmod 4$，则选手一获胜，而若 $k\equiv 1 \pmod 4$，则选手二获胜。假设现在玩 $k+1$ 根火柴的游戏，$k\geqslant 4$。如果 $k+1\equiv 0 \pmod 4$，则选手一拿走 3 根火柴，给对手留下 $k-2$ 根火柴。由于 $k-2\equiv 1 \pmod 4$，根据归纳假设，此时这是个后手(即原来的选手一)获胜的游戏。同理，如果 $k+1\equiv 2 \pmod 4$，则选手一拿走 1 根火柴；如果 $k+1\equiv 3 \pmod 4$，则选手一拿走 2 根火柴。最后，如果 $k+1\equiv 1 \pmod 4$，则选手一必定给对手留下 k、$k-1$ 或$k-2$根火柴。由于 $k\equiv 0 \pmod 4$，$k-1\equiv 3 \pmod 4$，$k-2\equiv 2 \pmod 4$，根据归纳假设，此时这是个先手(即原来的选手二)获胜的游戏。

37. *基础步骤*：对 $n=1$ 来说，左边只是 1/1，它是 1。对 $n=2$ 来说，存在三个非空子集合$\{1\}$、$\{2\}$和$\{1, 2\}$，所以左边是 $1/1+1/2+1/1\cdot 2=2$。

归纳步骤：假设对 k 来说命题为真。前 $k+1$ 个正整数的集合有许多非空子集合，但它们分成三个类别：前 k 个正整数的非空子集合，前 k 个正整数的非空子集合加上 $k+1$，或仅仅$\{k+1\}$。根据归纳假设，第一个类别之和为 k。对第二个类别来说，可以从和的每项里分解出 $1/(k+1)$来，根据归纳假设，剩下的和恰是 k，所以这个部分的和是 $k/(k+1)$。最后，第三个类别只产生出 $1/(k+1)$。因此，总和为 $k+k/(k+1)+1/(k+1)=k+1$。

39. 用记号(i, j)表示第 i 行第 j 列的棋盘格，对 $i+j$ 进行归纳来证明马能到达所有格子。

基础步骤：对于 $i+j\leqslant 2$ 的情形，有 6 种基本情形。马出发时已经在$(0, 0)$格了，所以不用走就到了这个格子。为了到达$(1, 0)$，马走$(0, 0)\to(2, 1)\to(0, 2)\to(1, 0)$。同理，为了到达$(0, 1)$，马走$(0, 0)\to(1, 2)\to(2, 0)\to(0, 1)$。注意在这个过程中马已经到了$(2, 0)$和$(0, 2)$。最后一种基本情形是$(0, 0)\to(1, 2)\to(2, 0)\to(0, 1)\to(2, 2)\to(0, 3)\to(1, 1)$。

归纳步骤：假设归纳假设，即马可以到达 $i+j=k$ 的任意(i, j)格子，其中 k 是大于 1 的整数。必须证明马可以到达 $i+j=k+1$ 的任意(i, j)格子。由于 $k+1\geqslant 3$，i 和 j 中至少有一个不小于 2，如果 $i\geqslant 2$，则根据归纳假设，有走法到达$(i-2, j+1)$，因为 $i-2+j+1=i+j-1=k$；由此再走一步就到(i, j)。如果 $j\geqslant 2$，则是类似的。

41. 基础情形 $n=0$ 和 $n=1$ 都为真，因为 x^0 的导数是 0 而且 $x^1=x$ 的导数是 1。利用乘法规则，归纳假设以及基础情形，就证明$\frac{d}{dx}x^{k+1}=\frac{d}{dx}(x\cdot x^k)=x\cdot\frac{d}{dx}x^k+x^k\frac{d}{dx}x=x\cdot kx^{k-1}+x^k\cdot 1=kx^k+x^k=(k+1)x^k$。

43. 设 $P(n)$是命题：$\mathbf{AB}^n=\mathbf{B}^n\mathbf{A}$。$P(1)$为真，因为 $\mathbf{AB}=\mathbf{BA}$。现在假设 $P(n)$为真。则 $\mathbf{AB}^{n+1}=\mathbf{AB}^n\mathbf{B}=\mathbf{B}^n\mathbf{AB}=\mathbf{B}^n\mathbf{BA}=\mathbf{B}^{n+1}\mathbf{A}$。所以 $P(n+1)$为真。

45. 设 $P(n)$是“$(A_1\cup A_2\cup\cdots\cup A_n)\cap B=(A_1\cap B)\cup(A_2\cap B)\cup\cdots\cup(A_n\cap B)$”。

基础步骤：$P(1)$平凡地为真。

归纳步骤：假设 $P(k)$为真，则$(A_1\cup A_2\cup\cdots\cup A_{k+1})\cap B=[(A_1\cup A_2\cup\cdots\cup A_k)\cup A_{k+1}]\cap B=[(A_1\cup A_2\cup\cdots\cup A_k)\cap B]\cup(A_{k+1}\cap B)=[(A_1\cap B)\cup(A_2\cap B)\cup\cdots\cup(A_k\cap B)]\cup(A_{k+1}\cap B)=(A_1\cap B)\cup(A_2\cap B)\cup\cdots\cup(A_k\cap B)\cup(A_{k+1}\cap B)$。

47. 设 $P(n)$是

$$“\overline{\bigcup_{k=1}^{n}A_k}=\bigcap_{k=1}^{n}\overline{A_k}”$$

基础步骤：$P(1)$平凡地为真。

归纳步骤：假设 $P(k)$为真，则

$$\overline{\bigcup_{k=1}^{k+1} A_k} = \overline{\bigcup_{k=1}^{k} A_k \cup A_{k+1}} = \overline{\bigcup_{k=1}^{k} A_k} \cap \overline{A_{k+1}} = \bigcap_{k=1}^{k} \overline{A_k} \cap \overline{A_{k+1}} = \bigcap_{k=1}^{k+1} \overline{A_k}$$

49. 设 $P(n)$是“$[(p_1 \to p_2) \land (p_2 \to p_3) \land \cdots \land (p_{n-1} \to p_n)] \to [(p_1 \land p_2 \land \cdots \land p_{n-1}) \to p_n]$”。

基础步骤：$P(2)$真，因为$(p_1 \to p_2) \to (p_1 \to p_2)$是重言式。

归纳步骤：假设$P(k)$为真。为了证明$[(p_1 \to p_2) \land (p_2 \to p_3) \land \cdots \land (p_{k-1} \to p_k) \land (p_k \to p_{k+1})] \to [(p_1 \land p_2 \land \cdots \land p_{k-1} \land p_k) \to p_{k+1}]$是重言式，假设这个蕴含式的前件为真。因为前件和$P(k)$都为真，所以$(p_1 \land p_2 \land \cdots \land p_{k-1}) \to p_k$ 为真。因为这个式子为真，并且因为 $p_k \to p_{k+1}$（它是假设的一部分），所以根据假言三段论，$(p_1 \land p_2 \land \cdots \land p_{k-1}) \to p_{k+1}$为真。从这个式子得出较弱的命题$(p_1 \land p_2 \land \cdots \land p_{k-1} \land p_k) \to p_{k+1}$。

51. 若 $n+1=2$，则这两个集合不重叠。事实上，蕴含式 $P(1) \to P(2)$为假。

53. 错误在于从基础情形 $n=0$ 过渡到下一情形 $n=1$，不能把 1 写成两个更小的自然数之和。

55. 假设良序性成立。假设 $P(1)$为真和对每一个正整数 $n \geqslant 1$ 来说，$(P(1) \land P(2) \land \cdots \land P(n)) \to P(n+1)$为真。设 S 是 $P(n)$为假的整数 n 的集合，将证明 $S=\varnothing$。假设 $S \neq \varnothing$。则根据良序性，在 S 里存在最小整数 m。m 不能是 1，因为 $P(1)$为真。因为 $n=m$ 是使得 $P(n)$为假的最小整数，所以 $P(1)$，$P(2)$，…，$P(m-1)$都为真，并且 $m-1 \geqslant 1$。因为$(P(1) \land P(2) \land \cdots \land P(m-1)) \to P(m)$为真，所以 $P(m)$必然为真，这是矛盾。因此 $S=\varnothing$。

57. 设 $P(n)$是“$H_{2^n} \leqslant 1+n$”。

基础步骤：$P(0)$为真，因为 $H_{2^0}=H_1=1 \leqslant 1+0$。

归纳步骤：假设 $H_{2^k} \leqslant 1+k$。则

$$H_{2^{k+1}} = H_{2^k} + \sum_{j=2^k+1}^{2^{k+1}} \frac{1}{j} \leqslant 1+k+2^k\left(\frac{1}{2^{k+1}}\right) < 1+k+1 = 1+(k+1)$$

59. 设 $P(n)$是“$1/\sqrt{1}+1/\sqrt{2}+1/\sqrt{3}+\cdots+1/\sqrt{n} > 2(\sqrt{n+1}-1)$”。

基础步骤：$P(1)$为真，因为 $1>2(\sqrt{2}-1)$。

归纳步骤：假设 $P(k)$为真。则 $1/\sqrt{1}+1/\sqrt{2}+1/\sqrt{3}+\cdots+1/\sqrt{k}+1/\sqrt{k+1} > 2(\sqrt{k+1}-1)+1/\sqrt{k+1}$。若证明 $2(\sqrt{k+1}-1)+1/\sqrt{k+1} > 2(\sqrt{k+2}-1)$，则得出 $P(k+1)$为真。这个不等式等价于 $2(\sqrt{k+2}-\sqrt{k+1}) < 1/\sqrt{k+1}$，它等价于 $2(\sqrt{k+2}-\sqrt{k+1})(\sqrt{k+2}+\sqrt{k+1}) < \sqrt{k+1}/\sqrt{k+1}+\sqrt{k+2}/\sqrt{k+1}$。这等价于 $2<1+\sqrt{k+2}/\sqrt{k+1}$，它显然为真。

61. 将首先证明当 n 是 2 的幂时的结果，即若 $n=2^k$，$k=1, 2, \cdots$。设 $P(k)$是命题：$A \geqslant G$，其中 A 和 G 是一组 $n=2^k$ 个正实数的算术和几何平均值。

基础步骤：$k=1$ 并且 $n=2^1=2$。注意$(\sqrt{a_1}-\sqrt{a_2})^2 \geqslant 0$。展开这个式子证明 $a_1-2\sqrt{a_1a_2}+a_2 \geqslant 0$，即$(a_1+a_2)/2 \geqslant (a_1a_2)^{1/2}$。

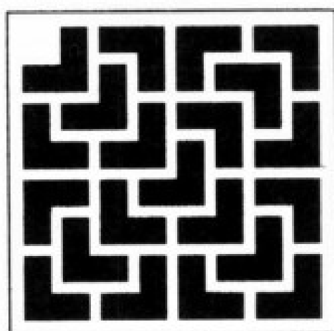
归纳步骤：假设 $P(k)$为真，并且 $n=2^k$。将证明 $P(k+1)$为真。我们有 $2^{k+1}=2n$，现在$(a_1+a_2+\cdots+a_{2n})/(2n)=((a_1+a_2+\cdots+a_n)/n+(a_{n+1}+a_{n+2}+\cdots+a_{2n})/n)/2$，同理，$(a_1a_2\cdots a_{2n})^{1/(2n)}=[(a_1a_2\cdots a_n)^{1/n}(a_{n+1}a_{n+2}\cdots a_{2n})^{1/n}]^{1/2}$。为了简化记号，设$A(x, y, \cdots)$和$G(x, y, \cdots)$分别表示 x，y，…的算术平均值和几何平均值。另外，若 $x \leqslant x'$，$y \leqslant y'$，依次类推，则 $A(x, y, \cdots) \leqslant A(x', y', \cdots)$而且 $G(x, y, \cdots) \leqslant G(x', y', \cdots)$。因此，$A(a_1, a_2, \cdots, a_{2n})=A(A(a_1, a_2, \cdots, a_n), A(a_{n+1}, a_{n+2}, \cdots, a_{2n})) \geqslant A(G(a_1, a_2, \cdots, a_n), G(a_{n+1}, a_{n+2}, \cdots, a_{2n})) \geqslant G(G(a_1, a_2, \cdots, a_n), G(a_{n+1}, a_{n+2}, \cdots, a_{2n}))=G(a_1, a_2, \cdots, a_{2n})$。这样就完成了对 2 的幂的情形的证明。现在若 n 不是 2 的幂，则设 m 是 2 的下一个更高的幂，并且设 a_{n+1}，…，a_m 都等于 $A(a_1, a_2, \cdots, a_n)=\bar{a}$。于是就有$((a_1a_2\cdots a_n)\bar{a}^{m-n})^{1/m} \leqslant A(a_1, a_2, \cdots, a_m)$，因为 m 是 2 的幂。因为 $A(a_1, a_2, \cdots, a_m)=\bar{a}$，所以$(a_1a_2\cdots a_k)^{1/m}\bar{a}^{1-n/m} \leqslant \bar{a}^{n/m}$。两边都求第 m/k 次幂就得出 $G(a_1, a_2, \cdots, a_n) \leqslant A(a_1, a_2, \cdots, a_n)$。

63. 当 $n=1$ 时，对基础情形来说，没有什么需要证明的。现在假设归纳假设成立。假设 $p \mid a_1a_2\cdots a_ka_{k+1}$。注意 $\gcd(p, a_1, a_2, \cdots, a_k)=1$ 或 p。若它是 1，则根据 2.6 节引理 1，有 $p \mid a_{k+1}$。若它是 p，则 $p \mid a_1a_2\cdots a_k$，所以根据归纳假设，对某个 $i\leqslant k$ 来说有 $p \mid a_i$。这样就完成了证明。

65. 根据提示，首先要证明，如果 $w^2+x^2+y^2+z^2=2wxyz$，则所有变元都是偶数。显然有 0、2 或 4 个变元是偶数，因为方程右边是偶数。假设所有变元都是奇数，则左边是 $1+1+1+1\equiv 0 \pmod 4$（参见 3.1 节练习 2）。但是在这种情形下右边模 4 同余 2，矛盾。下面假设变元是 2 奇 2 偶。于是左边模 4 为 2，而右边模 4 为 0，还是矛盾。所以就得出所有变元都是偶数，代入 $w=2a$，$x=2b$，$y=2c$，$z=2d$ 并化简，得到 $a^2+b^2+c^2+d^2=8abcd$。模 8 重复上述论证，得出 $a=2e$，$b=2f$，$c=2g$，$d=2h$，因此 $e^2+f^2+g^2+h^2=32efgh$，可以永远这样进行下去。所以每个原始变元都含有任意大的 2 的幂作为因子，而这是不可能的。

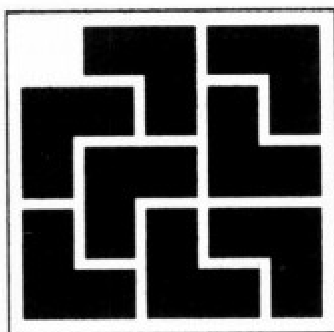
67. 设 $P(n)$ 是命题：若 x_1，x_2，$\cdots$，x_n 是 n 个不同的实数，则无论在乘积里如何插入括号，都需要用 $n-1$ 次乘法来求出这些数的乘积。将用数学归纳法第二原理来证明 $P(n)$ 为真。基础情形 $P(1)$ 为真，因为需要 $1-1=0$ 次乘法来求出 x_1 的乘积，即只有一个因子的乘积。假设对 $1\leqslant k\leqslant n$ 来说 $P(k)$ 为真。求出 $n+1$ 个不同实数 x_1，x_2，$\cdots$，x_n，x_{n+1} 的乘积所用的最后一次乘法，是对某个 k 来说这些数的前 k 个数的乘积与后 $n+1-k$ 个数的乘积之间的乘法。根据归纳假设，需要用 $k-1$ 次乘法来求出这些数中 k 个数的乘积，无论在这些数的乘积里如何插入括号，而且需要用 $n-k$ 次乘法来求出这些数中其余 $n+1-k$ 个数的乘积，无论在这些数的乘积里如何插入括号。因为需要另外一次乘法来求出所有这些 $n+1$ 个数的乘积，所以使用的乘法总次数等于 $(k-1)+(n-k)+1=n$。因此 $P(n+1)$ 为真。证毕。

69.



71. 设 $P(n)$ 是命题：每一个去掉了一个 $1\times1\times1$ 立方体的 $2^n\times2^n\times2^n$ 棋盘，都可以用每个去掉了一个 $1\times1\times1$ 立方体的 $2\times2\times2$ 立方体砖块来铺满。基础步骤 $P(1)$ 成立，因为一个砖块与所需要铺满的立体是重合的。现假设 $P(k)$ 成立。现在考虑去掉了一个 $1\times1\times1$ 立方体的 $2^{k+1}\times2^{k+1}\times2^{k+1}$ 立方体目标。用经过目标中心并且与目标表面平行的平面来把这个目标分成八块。所去掉的 $1\times1\times1$ 块出现在这八块当中的一块里。现在放一块砖，让它的中心是在这个大目标的中心上，使得它所去掉的 $1\times1\times1$ 立方体是在大目标去掉一个 $1\times1\times1$ 立方体的象限里。这样就产生出八个 $2^k\times2^k\times2^k$ 立方体，每个都去掉一个 $1\times1\times1$ 立方体。根据归纳假设，可以用砖块铺满这八个目标当中的每一个。把这些铺砖系统放在一起就产生出所需要的铺砖系统。

73.



75. 假定 $a=dq+r=dq'+r'$ 使得 $0\leqslant r<d$ 并且 $0\leqslant r'<d$，则 $d(q-q')=r'-r$。由此得出 d 整除 $r'-r$。因为 $-d<r'-r<d$，所以 $r'-r=0$。因此 $r'=r$。由此得出 $q=q'$。

77. 这是自引用所引起的悖论，答案显然是“否”。只有有穷多个英语单词，所以只有有穷多个包含 15 个或更少单词的字符串；因此只有有穷多个正整数可以如此描述，而不是所有的正整数。

3.4 节

1. a) $f(1)=3$，$f(2)=5$，$f(3)=7$，$f(4)=9$
   b) $f(1)=3$，$f(2)=9$，$f(3)=27$，$f(4)=81$
   c) $f(1)=2$，$f(2)=4$，$f(3)=16$，$f(4)=65\ 536$
   d) $f(1)=3$，$f(2)=13$，$f(3)=183$，$f(4)=33\ 673$
3. a) $f(2)=-1$，$f(3)=5$，$f(4)=2$，$f(5)=17$
   b) $f(2)=-4$，$f(3)=32$，$f(4)=-4096$，$f(5)=536\ 870\ 912$
   c) $f(2)=8$，$f(3)=176$，$f(4)=92\ 672$，$f(5)=25\ 764\ 174\ 848$
   d) $f(2)=-\frac{1}{2}$，$f(3)=-4$，$f(4)=\frac{1}{8}$，$f(5)=-32$
5. a) 无效
   b) $f(n)=1-n$。
   基础步骤：$f(0)=1=1-0$。
   归纳步骤：如果 $f(k)=1-k$，则 $f(k+1)=f(k)-1=1-k-1=1-(k+1)$。
   c) 若 $n>0$ 则 $f(n)=4-n$，且 $f(0)=2$。
   基础步骤：$f(0)=2$ 且 $f(1)=3=4-1$。
   归纳步骤：$(k\geqslant1)$：$f(k+1)=f(k)-1=(4-k)-1=4-(k+1)$。
   d) $f(n)=2^{\lfloor(n+1)/2\rfloor}$。
   基础步骤：$f(0)=1=2^{\lfloor(0+1)/2\rfloor}$ 且 $f(1)=2=2^{\lfloor(1+1)/2\rfloor}$。
   归纳步骤：$(k\geqslant1)$：$f(k+1)=2f(k-1)=2\cdot2^{\lfloor k/2\rfloor}=2^{\lfloor k/2\rfloor+1}=2^{\lfloor((k+1)+1)/2\rfloor}$。
   e) $f(n)=3^n$。
   基础步骤：显然。
   归纳步骤：对奇数 n，$f(n)=3f(n-1)=3\cdot3^{n-1}=3^n$。而对偶数 $n>1$，$f(n)=9f(n-2)=9\cdot3^{n-2}=3^n$。
7. 有多种可能的正确答案，将给出相对简单的答案。
   a) 对 $n\geqslant1$ 来说 $a_{n+1}=a_n+6$ 并且 $a_1=6$
   b) 对 $n\geqslant1$ 来说 $a_{n+1}=a_n+2$ 并且 $a_1=3$
   c) 对 $n\geqslant1$ 来说 $a_{n+1}=10a_n$ 并且 $a_1=10$
   d) 对 $n\geqslant1$ 来说 $a_{n+1}=a_n$ 并且 $a_1=5$
9. $F(0)=0$，对 $n\geqslant1$ 来说 $F(n)=F(n-1)+n$
11. $P_m(0)=0$，$P_m(n+1)=P_m(n)+m$
13. 设 $P(n)$ 是“$f_1+f_3+\cdots+f_{2n-1}=f_{2n}$”。
    基础步骤：$P(1)$ 为真，因为 $f_1=1=f_2$。
    归纳步骤：假设 $P(k)$ 为真。则 $f_1+f_3+\cdots+f_{2n-1}+f_{2n+1}=f_{2n}+f_{2n-1}=f_{2n+2}+f_{2(n+1)}$。
15. **基础步骤**：$f_0f_1+f_1f_2=0\cdot1+1\cdot1=1^2=f_2^2$。
    归纳步骤：假设 $f_0f_1+f_1f_2+\cdots+f_{2k-1}f_{2k}=f_{2n}^2$。则 $f_0f_1+f_1f_2+\cdots+f_{2n-1}f_{2n}+f_{2n}f_{2n+1}+f_{2n+1}f_{2n+2}=f_{2n}^2+f_{2n}f_{2n+1}+f_{2n+1}f_{2n+2}=f_{2n}(f_{2n}+f_{2n+1})+f_{2n+1}f_{2n+2}=f_{2n}f_{2n+2}+f_{2n+1}f_{2n+2}=(f_{2n}+f_{2n+1})f_{2n+2}=f_{2n+2}^2$。
17. 欧几里得算法为求出 $\gcd(f_{n+1},f_n)$ 而使用的除法次数，对 $n=0$ 来说是 0，对 $n=1$ 来说是 1，而对 $n\geqslant2$ 来说是 $n-1$。为了对 $n\geqslant2$ 证明这个结果，用数学归纳法。对 $n=2$ 来说，一次除法说明 $\gcd(f_3,f_2)=\gcd(2,1)=\gcd(1,0)=1$。现在假设用 $k-1$ 次除法来求出 $\gcd(f_{n+1},f_n)$。为了求出 $\gcd(f_{n+2},f_{n+1})$，首先让 f_{n+2} 除以 f_{n+1} 来得出 $f_{n+2}=1\cdot f_{n+1}+f_n$。在一次除法之后有 $\gcd(f_{n+2},f_{n+1})=\gcd(f_{n+1},f_n)$。根据归纳假设，所以恰好需要另外 $n-1$ 次除法。这说明需要 n 次除法来求出 $\gcd(f_{n+2},f_{n+1})$，证毕。
19. $|A|=-1$。因此 $|A^n|=(-1)^n$。所以 $f_{n+1}f_{n-1}-f_n^2=(-1)^n$。

21. a)用归纳法来证明。

基础步骤：对 $n=1$ 来说，$\max(-a_1)=-a_1=-\min(a_1)$。对 $n=2$ 来说，有两种情形。若 $a_2\geqslant a_1$，则 $-a_1\geqslant -a_2$，所以 $\max(-a_1,-a_2)=-a_1=-\min(a_1,a_2)$。若 $a_2<a_1$，则 $-a_1\leqslant -a_2$，所以 $\max(-a_1,-a_2)=-a_2=-\min(a_1,a_2)$。

归纳步骤：假设对满足 $k\geqslant 2$ 的 n 来说命题为真。则 $\max(-a_1,-a_2,\cdots,-a_n,-a_{n+1})=\max(\max(-a_1,-a_2,\cdots,-a_n),-a_{n+1})=\max(-\min(a_1,a_2,\cdots,a_n),-a_{n+1})=-\min(\min(-a_1,-a_2,\cdots,-a_n),a_{n+1})=-\min(a_1,\cdots,a_{n+1})$。

b)用数学归纳法来证明。对 $n=1$ 来说，结果是恒等式 $a_1+b_1=a_1+b_1$。对 $n=2$ 来说，首先考虑 $a_1+b_1\geqslant a_2+b_2$ 的情形。则 $\max(a_1+b_1,a_2+b_2)=a_1+b_1$。另外注意 $a_1\leqslant\max(a_1,a_2)$ 和 $b_1\leqslant\max(b_1,b_2)$，所以 $a_1+b_1\leqslant\max(a_1,a_2)+\max(b_1,b_2)$。因此 $\max(a_1+b_1,a_2+b_2)=a_1+b_1\leqslant\max(a_1,a_2)+\max(b_1,b_2)$。$a_1+b_1<a_2+b_2$ 的情形是类似的。对归纳步骤来说，假设对 k 来说结果为真。则 $\max(a_1+b_1,a_2+b_2,\cdots,a_n+b_n,a_{n+1}+b_{n+1})=\max(\max(a_1+b_1,a_2+b_2,\cdots,a_n+b_n),a_{n+1}+b_{n+1})\leqslant\max(\max(a_1,a_2,\cdots,a_n),\max(b_1,b_2,\cdots,b_n),a_{n+1}+b_{n+1})\leqslant\max(\max(a_1,a_2,\cdots,a_n),a_{n+1})+\max(\max(b_1,b_2,\cdots,b_n),b_{n+1})=\max(a_1,a_2,\cdots,a_n,a_{n+1})+\max(b_1,b_2,\cdots,b_n,b_{n+1})$。

c)与 b)一样，但是把"max"的每一次出现都换成"min"，并且反转每个不等式的方向。

23. $5\in S$，并且若 x，$y\in S$，则 $x+y\in S$。

25. a)$0\in S$，并且若 $x\in S$，则 $x+2\in S$ 和 $x-2\in S$。

b)$2\in S$，并且若 $x\in S$，则 $x+3\in S$。

c)$1\in S$，$2\in S$，$3\in S$，$4\in S$，并且若 $x\in S$，则 $x+5\in S$。

27. a)(0，1)，(1，1)，(2，1)；　(0，2)，(1，2)，(2，2)，(3，2)，(4，2)；　(0，3)，(1，3)，(2，3)，(3，3)，(4，3)，(5，3)，(6，3)；　(0，4)，(1，4)，(2，4)，(3，4)，(4，4)，(5，4)，(6，4)，(7，4)，(8，4)。

b)含 $P(n)$ 是命题：如果通过 n 次应用递归步骤得到 $(a,b)\in S$，则 $a\leqslant 2b$。

基础步骤：$P(0)$ 为真，因为不用递归步骤能得到的唯一 S 元素是 (0，0)，而确实 $0\leqslant 2\cdot 0$。

归纳步骤：假设如果通过不超过 k 次应用递归步骤得到 $(a,b)\in S$，则 $a\leqslant 2b$。考虑通过 $k+1$ 次应用递归步骤得到的一个元素。由于最后一步递归步骤应用到的元素 (a,b)，一定是通过较少次数应用递归步骤得到的，所以 $a\leqslant 2b$。两边分别加上 $0\leqslant 2$，$1\leqslant 2$，$2\leqslant 2$，就得到 $a\leqslant 2(b+1)$，$a+1\leqslant 2(b+1)$，$a+2\leqslant 2(b+1)$，这正是所需要的。

c)命题对基础步骤成立，因为 $0\leqslant 0$。假设命题对 (a,b) 成立，则命题对从 (a,b) 应用递归步骤得到的元素也成立，因为给 $a\leqslant 2b$ 分别加上 $0\leqslant 2$，$1\leqslant 2$，$2\leqslant 2$，就得到 $a\leqslant 2(b+1)$，$a+1\leqslant 2(b+1)$，$a+2\leqslant 2(b+1)$。

29. a)这样定义 S：$(1,1)\in S$；如果 $(a,b)\in S$，则 $(a+2,b)\in S$、$(a,b+2)\in S$ 且 $(a+1,b+1)\in S$。所有进入 S 的元素都满足条件，因为 (1，1) 坐标和是偶数，而如果 (a,b) 坐标和是偶数，则 $(a+2,b)$、$(a,b+2)$ 和 $(a+1,b+1)$ 也是这样。反之，对坐标和进行归纳来证明，如果 (a,b) 坐标和是偶数，则 $(a,b)\in S$。如果和是 2，则 $(a,b)=(1,1)$，基础步骤让 (a,b) 进入 S。否则和至少是 4，$(a-2,b)$、$(a,b-2)$ 和 $(a-1,b-1)$ 中至少有一个都是正整数坐标并且其和为小于 $a+b$ 的偶数，因此一定属于 S。于是再应用一次递归步骤就说明 $(a,b)\in S$。

b)这样定义 S：(1，1)、(1，2) 和 (2，1) 属于 S；如果 $(a,b)\in S$，则 $(a+2,b)$ 和 $(a,b+2)$ 都属于 S。要证明这个定义正确，首先注意 (1，1)、(1，2) 和 (2，1) 都有一个奇数坐标，而如果 (a,b) 有一个奇数坐标，则 $(a+2,b)$ 和 $(a,b+2)$ 也是这样。反之，对坐标和进行归纳来证明，如果 (a,b) 有一个奇数坐标，则 $(a,b)\in S$。如果 $(a,b)=(1,1)$ 或 $(a,b)=(1,2)$ 或 $(a,b)=(2,1)$，则基础步骤让 (a,b) 进入 S。否则 a 或 b 至少是 3，所以 $(a-2,b)$ 和 $(a,b-2)$ 中至少有一个都是正整数坐标并且其和小于 $a+b$，因此一定属于 S。于是再应用一次递归步骤就说明 $(a,b)\in S$。

c)$(1,6)\in S$ 且 $(2,3)\in S$；如果 $(a,b)\in S$，则 $(a+2,b)\in S$ 且 $(a,b+6)\in S$。要证明这个定义正确，

首先注意(1, 6)和(2, 3)满足条件，而且如果(a, b)满足条件，则$(a+2, b)$和$(a, b+6)$也是这样。反之，对坐标和进行归纳来证明，如果(a, b)满足条件，则$(a, b)\in S$。对于和为5或7的情形，唯一的点是：(1, 6)，基础步骤让其进入S；(2, 3)，基础步骤让其进入S；以及$(4, 3)=(2+2, 3)$，一次应用递归定义就让其进入S。对于和大于7的情形，要么$a\geqslant 3$，要么$a\leqslant 2$且$b\geqslant 9$，在这种情形下，$(a-2, b)$或$(a, b-6)$必定有和小于$a+b$的正整数坐标并且满足属于S的条件。于是再应用一次递归步骤就说明$(a, b)\in S$。

31. 若x是集合或是表示集合的变元，则x是合式公式。若x和y都是合式公式，则$\bar{x}$、$(x\cup y)$、$(x\cap y)$和$(x-y)$都是合式公式。

33. a)如果$x\in D=\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$，则$m(x)=x$；如果$s=tx$，其中$t\in D^*$且$x\in D$，则$m(s)=\min(m(s), x)$。

    b)令$t=wx$，其中$w\in D^*$且$x\in D$。如果$w=\lambda$，则根据m定义的归纳步骤和基础步骤，就有$m(st)=m(sx)=\min(m(s), x)=\min(m(s), m(x))$。否则，由$m$的定义，$m(st)=m((sw)x)=\min(m(sw), x)$。现在根据结构归纳法的归纳假设，$m(sw)=\min(m(sw), m(w))$，所以，根据min的含义，$m(st)=\min(\min(m(sw), m(w)), x)=\min(m(sw), \min(m(w), x))$。但是根据$m$定义的归纳步骤，$\min(m(w), x)=m(wx)=m(t)$。因此$\min(st)=\min(m(s), m(t))$。

35. $\lambda^R=\lambda$并且对$x\in\Sigma$来说$(ux)^R=xu^R$。

37. $w^0=\lambda$并且$w^{n+1}=ww^n$。

39. 当对某个非负整数n来说，该字符串是由n个0和后面的n个1所组成的时。

41. 设$P(i)$是“$l(w^i)=i\cdot l(w)$”。$P(0)$为真，因为$l(w^0)=0=0\cdot l(w)$。假设$P(i)$为真，则$l(w^{i+1})=l(ww^i)=l(w)+l(w^i)=l(w)+i\cdot l(w)=(i+1)\cdot l(w)$。

43. **基础步骤**：对于只含树根的满二叉树，结论为真，因为$n(T)=1$且$h(T)=0$，而$1\geqslant 2\cdot 0+1$。

    归纳步骤：假设$n(T_1)\geqslant 2h(T_1)+1$和$n(T_2)\geqslant 2h(T_2)+1$。根据$n(T)$和$h(T)$的递归定义，有$n(T)=1+n(T_1)+n(T_2)$和$h(T)=1+\max(h(T_1), h(T_2))$。因此$n(T)=1+n(T_1)+n(T_2)\geqslant 1+2h(T_1)+1+2h(T_2)+1\geqslant 1+2\cdot\max(h(T_1), h(T_2))+2\geqslant 1+2(\max(h(T_1), h(T_2))+1)=1+2h(T)$。

45. **基础步骤**：$a_{0,0}=0=0+0$。**归纳步骤**：假设在$N\times N$字典序下，只要(m', n')小于(m, n)就有$a_{m',n'}=m'+n'$。如果$n=0$，则$a_{m,n}=a_{m-1,n}+1=m-1+n+1=m+n$。如果$n>0$，则$a_{m,n}=a_{m,n-1}+1=m+n-1+1=m+n$。

47. a)$P_{m,m}=P_m$，因为在m的分拆里不可能使用比m大的数。

    b)因为只有一种方式来分拆1，即$1=1$，所以$P_{1,n}=1$。因为只有一种方式来把m分拆成1，所以$P_{m,1}=1$。当$n>m$时，得出$P_{m,n}=P(m, m)$，因为不可能使用比m大的数。$P_{m,m}=1+P_{m,m-1}$，因为当在分拆里允许使用m时，出现另外一个分拆，即$m=m$。若$m>n$，则$P_{m,n}=P_{m,n-1}+P_{m-n,n}$，因为一个把$m$分成不大于$n$的整数的分拆，要么不使用任何$n$，因此算在$P_{m,n-1}$里，要么使用一个$n$和一个对$m-n$的分拆，因此算在$P_{m-n,n}$里。

    c)$P_5=7$，$P_6=11$

49. 设$P(n)$是“$A(n, 2)=4$”。

    基础步骤：$P(1)$为真，因为$A(1, 2)=A(0, A(1, 1))=A(0, 2)=2\cdot 2=4$。

    归纳步骤：假设$P(n)$为真，即$A(n, 2)=4$，则$A(n+1, 2)=A(n, A(n+1, 1))=A(n, 2)=4$。

51. a)16　　b)65 536

53. 用双重归纳论证来证明更强的命题：当$k>l$时，$A(m, k)>A(m, l)$。

    基础步骤：当$m=0$时命题为真，因为$k>l$蕴含$A(0, k)=2k>2l=A(0, l)$。

    归纳步骤：假设对所有满足$x>y$的非负整数x和y来说都有$A(m, x)>A(m, y)$。将证明这个假设蕴含着：若$k>l$，则$A(m+1, k)>A(m+1, l)$。

    基础步骤：当$l=0$而且$k>0$时，$A(m+1, l)=0$，并且或者$A(m+1, k)=2$，或者$A(m+1, k)=A(m, A(m+1, k-1))$。若$m=0$，则它是$2A(1, k-1)=2^k$。若$m>0$，则根据归纳假设，它是大于0

的。在所有情形里 $A(m+1, k)>0$，事实上，$A(m+1, k)\geqslant 2$。若 $l=1$ 而且 $k>0$，$A(m+1, l)=2$ 并且或者 $A(m+1, k)=A(m, A(m+1, k-1))\geqslant 2$。因此根据归纳假设，$A(m, A(m+1, k-1))\geqslant A(m, 2)>A(m, 1)=2$。

归纳步骤：假设对所有 $r>s(s=0, 1, \cdots, l)$来说，都有 $A(m+1, r)>A(m+1, s)$。于是若 $k+1>l+1$，则得出 $A(m+1, k+1)=A(m, A(m+1, k))>A(m, A(m+1, l))=A(m+1, l+1)$。

55. 从练习 54 得出 $A(i, j)\geqslant A(i-1, j)\geqslant\cdots\geqslant A(0, j)=2j\geqslant j$。

57. 设 $P(n)$是"$F(n)$是良定义的"。则 $P(0)$为真，因为规定了 $F(0)$。假设对所有 $k<n$ 来说 $P(k)$为真。则在 n 上 $F(n)$是良定义的，因为 $F(n)$是利用 $F(0)$，$F(1)$，…，$F(n_1)$来给出的。所以对所有整数 n 来说 $P(n)$为真。

59. a) $F(1)$的值是二义性的。

    b) $F(2)$是无定义的，因为 $F(0)$是无定义的。

    c) $F(3)$是二义性的而且 $F(4)$是无定义的，因为 $F\left(\frac{4}{3}\right)$没有意义。

    d) $F(1)$的定义是二义性的，因为第二和第三个句子似乎都可以应用。

    e) 不能计算 $F(2)$，因为试图计算 $F(2)$会给出 $F(2)=1+F(F(1))=1+F(2)$。

61. a) 1　b) 2　c) 3　d) 3　e) 4　f) 4　g) 5

63. $f_0^*(n)=\lceil n/a\rceil$

65. 对 $n\geqslant 2$ 来说，$f_2^*(n)=\lceil\log\log n\rceil$，$f_2^*(1)=0$

3.5 节

1. **procedure** *mult*(n：正整数，x：整数)
   if $n=1$ **then** $mult(n, x):=x$
   else $mult(n, x):=x+mult(n-1, x)$

3. **procedure** *sum of odds*(n：正整数)
   if $n=1$ **then** $sum\ of\ odds(n):=1$
   else $sum\ of\ odds(n):=sum\ of\ odds(n-1)+2n-1$

5. **procedure** *smallest*($a_1, \cdots, a_n$：整数)
   if $n=1$ **then** $smallest(a_1, \cdots, a_n):=a_1$
   else $smallest(a_1, \cdots, a_n):=\min(smallest(a_1, \cdots, a_{n-1}), a_n)$

7. **procedure** *modfactorial*(n，m：正整数)
   if $n=1$ **then** $modfactorial(n, m):=1$
   else $modfactorial(n, m):=(n\cdot modfactorial(n-1, m))$ **mod** m

9. **procedure** gcd(a，b：非负整数)
   {假设 $a<b$ 是成立的}
   if $a=0$ **then** $\gcd(a, b):=b$
   else if $a=b-a$ **then** $\gcd(a, b):=a$
   else if $a<b-a$ **then** $\gcd(a, b):=\gcd(a, b-a)$
   else $\gcd(a, b):=\gcd(b-a, a)$

11. **procedure** *multiply*(x，y：非负整数)
    if $y=0$ **then** $multiply(x, y):=0$
    else if y 是偶数 **then**
      $multiply(x, y):=2x\cdot multiply(x, y/2)$
    else $multiply(x, y):=2x\cdot multiply(x, (y-1)/2)+x$

13. 如果 $n=1$，则 $nx=x$，算法正确返回 x。假设算法正确计算 kx。为了计算 $(k+1)x$，算法递归计算 $k+1-1=k$ 和 x 之积，然后加上 x。根据归纳假设，算法正确计算乘积，所以返回的答案是 $kx+x=(k+1)x$，这是对的。
15. 如果 $n=0$，则程序第一行正确返回函数值。假设程序对 $n=k$ 正确计算。如果 $n=k+1$，则执行 **else** 语句，返回值是 a 乘以 $power(a, k)$。由于根据归纳假设，$power(a, k)$ 是 a^k，所以返回值是 $a \cdot a^k = a^{k+1}$，这是对的。
17. n 次乘法与 2^n 次乘法
19. $O(\log n)$ 与 n
21.
```
procedure a(n: 非负整数)
if n=0 then a(n) := 1
else if n=1 then a(n) := 2
else a(n) := a(n-1) * a(n-2)
```
23. 迭代
25.
```
procedure iterative(n: 非负整数)
if n=0 then z := 1
else if n=1 then z := 2
else
begin
  x := 1
  y := 2
  z := 3
  for i := 1 to n-2
  begin
      w := x+y+z
      x := y
      y := z
      z := w
  end
end
{z 是该序列的第 n 项}
```
27. 首先给出递归过程，然后给出迭代过程。
```
procedure r(n: 非负整数)
if n<3 then r(n) := 2n+1
else r(n) = r(n-1) · (r(n-2))^2 · (r(n-3))^3
procedure i(n: 非负整数)
if n=0 then z := 1
else if n=1 then z := 3
else
begin
  x := 1
  y := 3
  z := 5
  for i := 1 to n-2
  begin
    w := z * y^2 * x^3
```



```
      x := y
      y := z
      z := w
    end
  end
  {z 是该序列的第 n 项}
  迭代过程更有效。
```

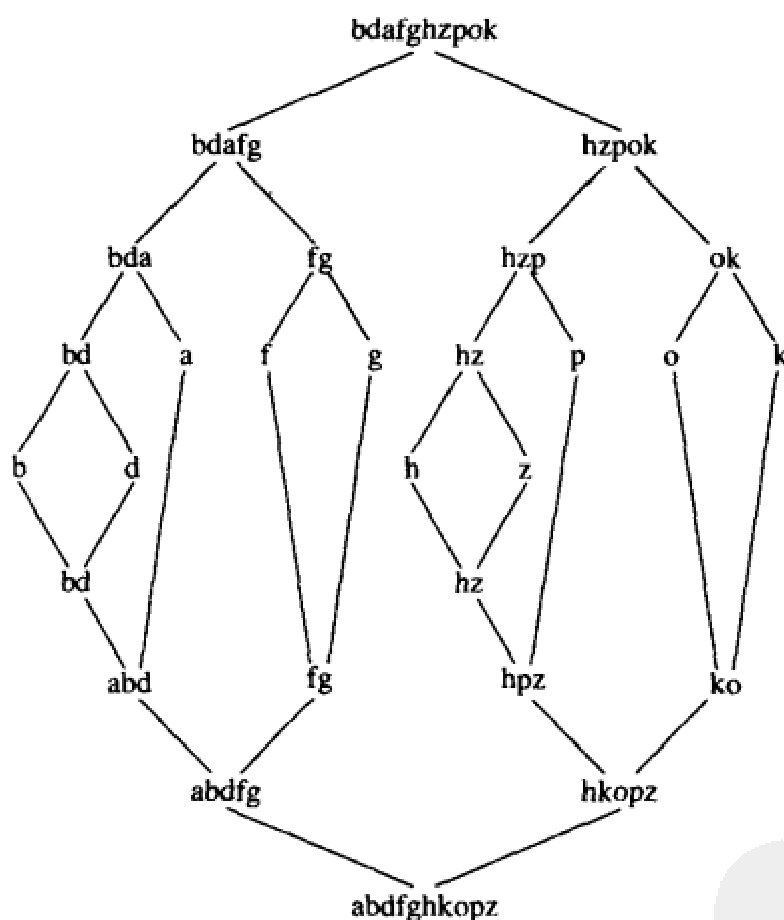
29. **procedure** *reverse*(w：位串)

```
n := length(w)
if n≤1 then reverse(w) := w
else reverse(w) := substr(w, n, n)reverse(substr(w, 1, n−1))
  {substr(w, a, b)是包括从第 a 个位置到第 b 个位置的 w 的子串}
```

31. **procedure** $A(m, n$：非负整数)

```
if m=0 then A(m, n) := 2n
else if n=0 then A(m, n) := 0
else if n=1 then A(m, n) := 2
else A(m, n) := A(m−1, A(m, n−1))
```

33.



35. 令这两个表分别为 1，2，…，$m-1$，$m+n-1$ 和 m，$m+1$，…，$m+n-2$，$m+n$。

37. 如果 $n=1$，则算法什么也不做，这是正确的，因为单个元素的表已经排序了。假设算法对于从 $n=1$ 到 $n=k$ 都是正确的。如果 $n=k+1$，则把表分成两个表 L_1 和 L_2。根据归纳假设，*mergesort* 正确排序这两个子表，而 *merge* 正确合并两个排序表成一个表，因为通过每次比较，就让 $L_1 \cup L_2$ 中还没有进入 L 的最小元素进入 L。

39. $O(n)$

41. 6

43. $O(n^2)$

3.6 节

1. 假设 $x=0$。这个程序段首先对 y 赋值 1 然后对 z 赋值 $x+y=0+1=1$。
3. 假设 $y=3$。这个程序段首先对 x 赋值 2 然后对 z 赋值 $x+y=2+3=5$。因为 $y=3>0$，所以它然后对 z 赋值$z+1=5+1=6$。
5. $(p \wedge condition)\{S_1\}q$

   $(p \wedge \neg condition1 \wedge condition2)\{S_2\}q$

   ⋮

   $(p \wedge \neg condition1 \wedge \neg condition2 \cdots \wedge \neg condition(n-1))\{S_n\}q$

   $\therefore p\{$**if** $condition1$ **then** S_1；

   else if $condition2$ **then** S_2；…；**else** $S_n\}q$
7. 将证明 p："$power=x^{i-1}$ 并且 $i\leqslant n+1$"是循环不变量。注意起初 p 为真，因为在循环开始前，$i=1$ 并且 $power=1=x^0=x^{1-1}$。其次必须证明在循环的一遍执行后，若 p 为真并且 $i\leqslant n$，则在循环的下一遍执行后，p 仍然为真。循环把 i 增加 1。因此，因为在这遍循环之前 $i\leqslant n$，所以在这遍循环之后 $i\leqslant n+1$。而且该循环对 $power$ 赋值 $power \cdot x$。根据归纳假设，可以看到把 $power$ 赋值为 $x^{i-1}\cdot x=x^i$。因此 p 仍然为真。另外，在该循环执行 n 遍之后，循环结束并且 $i=n+1$，因为在进入循环前把 i 赋值为 1，在每遍循环里把 i 增加 1，并且当 $i>n$ 时循环结束。所以，在结束时 $power=x^n$，正如所需。
9. 假设 p 是"m 和 n 都是整数"。于是若条件 $n<0$ 为真，则在执行 S_1 之后有 $a=-n=|n|$。若条件 $n<0$ 为假，则在执行 S_1 之后有 $a=n=|n|$。因此 $p\{S_1\}q$ 为真，其中 q 是 $p\wedge(a=|n|)$。因为 S_2 赋值 0 给 k 和 x，所以显然 $q\{S_2\}r$ 为真，其中 r 是 $q\wedge(k=0)\wedge(x=0)$。假设 r 为真。设 $P(k)$是"$x=mk$并且$k\leqslant a$"。可以证明对 S_3 里的循环来说 $P(k)$是循环不变量。$P(0)$为真，因为在进入循环前 $x=0=m\cdot 0$ 并且 $0\leqslant a$。现在假设 $P(k)$为真并且 $k<a$。则 $P(k+1)$为真，因为把 x 赋值为 $x+m=mk+m=m(k+1)$。当 $k=a$ 时循环结束，而在这个时候 $x=ma$。因此 $r\{S_3\}s$ 为真，其中 s 是"$a=|n|$ 并且 $x=ma$"。现在假设 s 为真。于是若 $n<0$，则 $a=-n$，所以 $x=-mn$。在这种情形里，S_4 给 $product$ 赋值 $-x=mn$。若 $n>0$,则 $x=ma=mn$，所以 S_4 给 $product$ 赋值 mn。因此 $s\{S_4\}t$ 为真。
11. 假设初始断言 p 为真。则因为 $p\{S\}q_0$ 为真，所以在执行程序段 S 后 q_0 为真。因为 $q_0\rightarrow q_1$ 为真，所以也得出在执行程序段 S 后 q_1 为真。因此 $p\{S\}q_1$ 为真。
13. 将用命题 p："$\gcd(a,b)=\gcd(x,y)$并且 $y\geqslant 0$"来作为循环不变量。注意在进入循环前 p 为真，因为利用初始断言，那时 $x=a$，$y=b$，且 y 是正整数。现在假设 p 为真并且 $y>0$，则循环将再度执行。在循环内部，x 和 y 分别换成 y 和 x **mod** y。根据 2.5 节的引理 1，$\gcd(x,y)=\gcd(y, x \textbf{ mod } y)$。因此，在循环执行后，$\gcd(x,y)$的值与循环执行前是一样的。另外，因为 y 是余数，所以它至少为 0。因此 p 仍然为真，所以它是循环不变量。另外，若循环结束，则 $y=0$。在这种情形里有 $\gcd(x,y)=x$，即终结断言。因此这个给出 y 作为输出的程序，正确地计算了 $\gcd(a,b)$。最后，可以证明循环必然结束，因为每次迭代导致 y 的值至少减少 1。因此，循环至多可以迭代 b 次。

补充练习

1. $2^{2^5}+1=4\,294\,967\,297=641\times 6\,700\,417$
3. 反证法：假设$\sqrt[3]{7}=p/q$，其中 p 和 q 是既约正整数。于是 $p^3=7q^3$。因此 7 是 p^3 的因子，所以也是 p 的因子(否则，p^3 的素因子分解根本不含 7)。对于某个正整数 k，写 $p=7k$。这样给出 $49k^3=q^3$。因此 7 是 q 的因子(与上述理由一样)。但是这与 p 和 q 没有非平凡公因子的事实相矛盾，因此$\sqrt[3]{7}$不是有理数。
5. 唯一可能的加项是 $1^4=1$，$2^4=16$，$3^4=81$，$4^4=256$ 和 $5^4=625$。显然两两相加都不等于 1000。
7. 令 $z=x-y$，所以 $x=y+z$。于是要证明的不等式变成 $|z|\geqslant|y+z|-|y|$，它等价于 $|y|+|z|\geqslant|y+z|$。但这是 1.5 节练习 35 证明过的不等式。

9. 令 $P(x)$ 如提示中那样。则 $P(x)$ 是多项式（其实是 $n-1$ 次的）；如果 $x=x_m$，则 $\prod_{i\neq j}(x-x_j)/(x_i-x_j)=0$，除非 $i=m$。因此 $P(x_m)=\prod_{j\neq m}y_m(x_m-x_j)/(x_m-x_j)=1\cdot y_m=y_m$。

11. 令 $P(n)$ 是"$1\cdot 2^0+2\cdot 2^1+3\cdot 2^2+\cdots+n\cdot 2^{n-1}=(n-1)\cdot 2^n+1$"。

基础步骤：$P(1)$ 为真，因为 $1\cdot 1=1=(1-1)2^1+1$。

归纳步骤：假设 $P(k)$ 为真，于是 $1\cdot 1+2\cdot 2+\cdots+k\cdot 2^{k-1}+(k+1)\cdot 2^k=(k-1)2^k+1+(k+1)2^k=2k\cdot 2^k+1=((k+1)-1)2^{k+1}+1$。

13. 令 $P(n)$ 是"$1/(1\cdot 4)+\cdots+1/[(3n-2)(3n+1)]=n/(3n+1)$"。

基础步骤：$P(1)$ 为真，因为 $1/(1\cdot 4)=1/4$。

归纳步骤：假设 $P(k)$ 为真，于是 $1/(1\cdot 4)+\cdots+1/[(3k-2)(3k+1)]+1/[(3k+1)(3k+4)]=k/(3k+1)+1/[(3k+1)(3k+4)]=[k(3k+4)+1]/[(3k+1)(3k+4)]=[(3k+1)(k+1)]/[(3k+1)(3k+4)]=(k+1)/(3k+4)$。

15. 令 $P(n)$ 是"$2^n>n^3$"。

基础步骤：$P(0)$ 为真，因为 $1024>1000$。

归纳步骤：假设 $P(k)$ 为真，于是 $(k+1)^3=k^3+3k^2+3k+1\leqslant k^3+9k^2\leqslant k^3+k^3=2k^3<2\cdot 2^k=2^{k+1}$。

17. 令 $P(n)$ 是"$a-b$ 是 a^n-b^n 的因子"。

基础步骤：$P(1)$ 显然为真。

归纳步骤：假设 $P(k)$ 为真，于是 $a^{k+1}-b^{k+1}=a^{k+1}-ab^k+ab^k-b^{k+1}=a(a^k-b^k)+b^k(a-b)$。则由于 $a-b$ 是 a^k-b^k 的因子并且 $a-b$ 是 $a-b$ 的因子，所以 $a-b$ 是 $a^{k+1}-b^{k+1}$ 的因子。

19. 令 $P(n)$ 是"$a+(a+d)+\cdots+(a+nd)=(n+1)(2a+nd)/2$"。

基础步骤：$P(1)$ 为真，因为 $a+(a+d)=2a+d=2(2a+d)/2$。

归纳步骤：假设 $P(k)$ 为真。于是 $a+(a+d)+\cdots+(a+kd)+(a+(k+1)d)=(k+1)(2a+kd)/2+a+(k+1)d=\frac{1}{2}[2ak+2a+k^2d+kd+2a+2kd+2d]=\frac{1}{2}[2ak+4a+k^2d+3kd+2d]=\frac{1}{2}(k+2)(2a+(k+1)d)$。

21. **基础步骤**：对 $n=1$ 为真，因为 $5/6=10/12$。

归纳步骤：假设对 $n=k$ 等式成立，考虑 $n=k+1$。于是 $\sum_{i=1}^{k+1}\frac{i+4}{i(i+1)(i+2)}=\sum_{i=1}^{k}\frac{i+4}{i(i+1)(i+2)}+\frac{k+5}{(k+1)(k+2)(k+3)}=\frac{k(3k+7)}{2(k+1)(k+2)}+\frac{(k+5)}{(k+1)(k+2)(k+3)}=$（根据归纳假设）$=\frac{1}{(k+1)(k+2)}\cdot\left(\frac{k(3k+7)}{2}+\frac{k+5}{k+3}\right)=\frac{1}{2(k+1)(k+2)(k+3)}\cdot(k(3k+7)(k+3)+2(k+5))=\frac{1}{2(k+1)(k+2)(k+3)}\cdot(3k^3+16k^2+23k+10)=\frac{1}{2(k+1)(k+2)(k+3)}\cdot(3k+10)(k+1)^2=\frac{1}{2(k+2)(k+3)}\cdot(3k+10)(k+1)=\frac{(k+1)(3(k+1)+7)}{2((k+1)+1)((k+1)+2)}$。

23. **基础步骤**：对 $n=1$ 命题成立，因为根据乘积求导法则，$g(x)=xe^x$ 的导数 $x\cdot e^x+e^x=(x+1)e^x$。

归纳步骤：假设对 $n=k$ 命题成立，即 k 重导数为 $g^{(k)}(x)=(x+k)e^x$。根据乘积求导法则对两边求导，给出 $g^{(k+1)}(x)=(x+k)e^x+e^x=(x+(k+1))e^x$，这正是所要的。

25. 将用数学归纳法第二原理来证明：若 $n\equiv 0(\mathrm{mod}\ 3)$，则 f_n 为偶数；否则 f_n 为奇数。基础步骤成立，因为 $f_0=0$ 是偶数而 $f_1=1$ 是奇数。现在假设如果 $k\leqslant n$，那么若 $k\equiv 0(\mathrm{mod}\ 3)$，则 f_k 为偶数，否则 f_k 为奇数。现在假设 $n+1\equiv 0(\mathrm{mod}\ 3)$，则 $f_{n+1}=f_n+f_{n-1}$ 是奇数，因为 f_n 是偶数而 f_{n-1} 是奇数。最后，若 $n+1\equiv 2(\mathrm{mod}\ 3)$，则 $f_{n+1}=f_n+f_{n-1}$ 是奇数，因为 f_n 是奇数而 f_{n-1} 是偶数。这样就完成了归纳证明。

27. 设 $P(n)$ 是命题：对每个非负整数 k 来说，有 $f_kf_n+f_{k+1}f_{n+1}=f_{n+k+1}$。基础步骤包括证明 $P(0)$ 和 $P(1)$ 都成立。$P(0)$ 为真，因为 $f_kf_0+f_{k+1}f_1=f_k\cdot 0+f_{k+1}\cdot 1=f_1$。因为 $f_kf_1+f_{k+1}f_2=f_k+f_{k+1}=$

f_{k+2}，所以 $P(1)$ 为真。现在假设 $P(j)$ 成立。则根据归纳假设和菲波那契数的递归定义，得出 $f_{k+1}f_{j+1}+f_{k+2}f_{j+2}=f_k(f_{j-1}+f_j)+f_{k+1}(f_j+f_{j+1})=(f_kf_{j-1}+f_{k+1}f_j)+(f_kf_j+f_{k+1}f_{j+1})=f_{j-1+k+1}+f_{j+k+1}=f_{j+k+2}$。这样就证明了 $P(j+1)$ 为真而且完成了证明。

29. 设 $P(n)$ 是命题：$l_0^2+l_1^2+\cdots+l_n^2=l_nl_{n+1}+2$。基础情形下 $P(0)$ 和 $P(1)$ 都成立，因为 $l_0^2=2^2=2\cdot1+2=l_0l_1+2$ 而且 $l_0^2+l_1^2=2^2+1^2=1\cdot3+2=l_1l_3+2$。现在假设 $P(k)$ 成立。则根据归纳假设，有 $l_0^2+l_1^2+\cdots+l_n^2+l_{n+1}^2=l_nl_{n+1}+2l_{n+1}^2=l_{n+1}(l_n+l_{n+1})+2=l_{n+1}l_{n+2}+2$。这样就证明了 $P(n+1)$ 成立而且完成了证明。

31. 设 $P(n)$ 是命题：对整数 n 来说，这个恒等式成立。基础情形 $P(1)$ 显然为真。假设 $P(k)$ 为真，则 $\cos(n+1)x+i\sin(n+1)x=\cos(nx+x)+i\sin(nx+x)=\cos nx\cos x-\sin nx\sin x+i(\sin nx\cos x+\cos nx\sin x)=\cos x(\cos nx+i\sin nx)(\cos x+i\sin x)=(\cos x+i\sin x)^n(\cos x+i\sin x)=(\cos x+i\sin x)^{n+1}$。所以 $P(n+1)$ 为真，证毕。

33. 把右边写成 $2^{n+1}(n^2-2n+3)-6$。对 $n=1$ 有 $2=4\cdot2-6$。假设对 $n=k$ 等式成立。考虑 $n=k+1$。于是 $\sum_{j=1}^{k+1}j^22^j=\sum_{j=1}^{k}j^22^j+(k+1)^22^{k+1}=2^{k+1}(k^2-2k+3)-6+(k^2+2k+1)2^{k+1}=$（根据归纳假设）$=2^{k+1}(2k^2+4)-6=2^{k+2}(k^2+2)-6=2^{k+2}(k+1)^2-2(k+1)+3)-6$。

35. 当 $n=1$ 时命题显然为真。假设对 $n=k$ 命题为真。考虑 $k+1$ 个人站成一队，女人打头男人收尾。如果第 k 个人是女人，则在队尾有一个女人站在一个男人前面。如果第 k 个人是男人，则队中前 k 个人就满足归纳假设的条件，所以又得出在队中某处有一个女人站在一个男人前面。

37. **基础步骤**：当 $n=1$ 时有一个圆周，可以把内部染成蓝色且把外部染成红色以满足条件。

归纳步骤：假设归纳假设，即如果有 k 个圆周，则这些区域可以 2- 着色，使得具有共同边界的区域都有不同颜色。考虑 $k+1$ 个圆周的情形。删除一个圆周，产生 k 个圆周的图形，利用归纳假设用所要求的方式对它着色。然后放回删除的圆周，并且改变其内部各个区域的颜色。得出的图形就满足条件，因为如果两个区域有共同边界，则要么这个边界属于新圆周，此时两边的区域原来是一个区域，现在里面的颜色与外面的颜色不同了，要么这个边界与新圆周无关，此时这两个区域染不同的颜色，因为在放回新圆周之前，这两个区域就染不同的颜色。

39. 如果 $n=1$，则方程是 $1\cdot1=1\cdot2/2$ 这是对的。假设对 $n=1$ 方程是对的，对 $n+1$ 考虑这个方程。则 $\sum_{j=1}^{n+1}(2j-1)\left(\sum_{k=j}^{n+1}\frac{1}{k}\right)=\sum_{j=1}^{n}(2j-1)\left(\sum_{k=j}^{n+1}\frac{1}{k}\right)+(2(n+1)-1)\cdot\frac{1}{n+1}=\sum_{j=1}^{n}(2j-1)\left(\frac{1}{n+1}+\sum_{k=j}^{n}\frac{1}{k}\right)+\frac{2n+1}{n+1}=\left(\frac{1}{n+1}\sum_{j=1}^{n}(2j-1)\right)+\left(\sum_{j=1}^{n}(2j-1)\sum_{k=j}^{n}\frac{1}{k}\right)+\frac{2n+1}{n+1}=\left(\frac{1}{n+1}\cdot n^2\right)+\frac{n(n+1)}{2}+\frac{2n+1}{n+1}=$（根据归纳假设）$=\frac{2n^2+n(n+1)^2+(4n+2)}{2(n+1)}=\frac{2(n+1)^2+n(n+1)^2}{2(n+1)}=\frac{(n+1)(n+2)}{2}$。

41. a) 92　b) 91　c) 91　d) 91　e) 91　f) 91

43. 基础步骤是不正确的，因为对所显示的和来说 $n\neq1$。

45. 设 $P(n)$ 是"若 n 个圆环中，每两个都有两个公共点，任何三个都没有公共点，则这些圆环把平面划分成 n^2-n+2 个区域"。

基础步骤：$P(1)$ 为真，因为一个圆环划分平面成 $2=1^2-1+2$ 个区域。

归纳步骤：假设 $P(k)$ 为真，即具有所规定性质的 k 个圆环划分平面成 k^2-k+2 个区域。假设添加第 $k+1$ 个圆环。这个圆环与其余 k 个圆环中的每个都相交于两点，所以这些交点形成 $2k$ 段新的圆弧，每段圆弧都分割一个老的区域。因此有 $2k$ 个区域被分割，这说明比以前多出 $2k$ 个区域。因此满足所规定性质的 $k+1$ 个圆环划分平面成 $k^2-k+2+2k=(k^2+2k+1)-(k+1)+2=(k+1)^2-(k+1)+2$ 个区域。

47. 假如 $\sqrt{2}$ 是有理数。则 $\sqrt{2}=a/b$，其中 a 和 b 都是正整数。所以集合 $S=\{n\sqrt{2}\mid n\in\mathbf{N}\}\cap\mathbf{N}$ 是正整数的非

空集合，因为 $b\sqrt{2}=a$ 属于 S。设 t 是 S 的最小元素，根据良序性，t 是存在的。于是对某个整数 s 来说有 $t=s\sqrt{2}$。就有 $t-s=s\sqrt{2}-s=s(\sqrt{2}-1)$，所以 $t-s$ 是正整数，因为$\sqrt{2}>1$。因此 $t-s$ 属于 S。这是一个矛盾，因为 $t-s=s\sqrt{2}-s<s$。因此$\sqrt{2}$是无理数。

49. 假设良序性为假。设 S 是没有最小元素的非负整数的非空集合。设 $P(n)$ 是命题"$i\notin S$，$i=0, 1, \cdots, n$"。$P(0)$为真，因为若 $0\in S$，则 S 有最小元素，即 0。现在假设 $P(n)$ 为真。因此 $0\notin S$，$1\notin S$，…，$n\notin S$。显然 $n+1$ 不可能属于 S，因为假如它属于 S，那么它就是最小元素。因此 $P(n+1)$为真。所以根据数学归纳法原理，对所有非负整数 n 来说，$n\in S$。因此 $S=\varnothing$，矛盾。

51. a)设 $d=\gcd(a_1, a_2, \cdots, a_n)$。则 d 是每个 a_i 的因子，所以 d 必然是 $\gcd(a_{n-1}, a_n)$的因子。因此 d 是 $a_1, a_2, \cdots, a_{n-2}$和 $\gcd(a_{n-1}, a_n)$的公因子。为了证明 d 是这些数的最大公因子，假设 c 是它们的公因子。则对 $i=1, 2, \cdots, n-2$ 来说，c 是 a_i 的因子，并且 c 是 $\gcd(a_{n-1}, a_n)$的因子，所以它是 a_{n-1} 和 a_n 的公因子。因此 c 是 $a_1, a_2, \cdots, a_{n-1}, a_n$ 的公因子。因此它是 $a_1, a_2, \cdots, a_n$ 的最大公因子 d 的因子。所以 d 是最大公因子，与所声称的一样。

    b)若 $n=2$，则应用欧几里得算法。否则，对 a_{n-1}和 a_n 应用欧几里得算法，得出 $d=\gcd(a_{n-1}, a_n)$，然后对 $a_1, a_2, \cdots, a_{n-2}$，$d$ 递归地应用这个算法。

53. $f(n)=n^2$。设 $P(n)$是"$f(n)=n^2$"。

    基础步骤：$P(1)$为真，因为 $f(1)=1=1^2$，这是从 f 的定义得出的。

    归纳步骤：假设 $f(n)=n^2$。则 $f(n+1)=f((n+1)-1)+2(n+1)-1=f(n)+2n+1=n^2+2n+1=(n+1)^2$。

55. a) λ，0，1，00，01，11，000，001，011，111，0000，0001，0011，0111，1111，00000，00001，00011，00111，01111，11111

    b) $S=\{\alpha\beta \mid \alpha$ 是由 m 个 0 所组成的串而 β 是由 n 个 1 所组成的串，$m\geqslant 0$，$n\geqslant 0\}$

57. 对 λ 应用第一条归纳步骤得出$()\in B$。对$()$应用第二条归纳步骤得出$()()\in B$。对$()()$应用第二条归纳步骤得出$(()())\in B$。由练习 60，$(()))$不属于 B，因为左右括号数不相等。

59. λ，()，(())，()()

61. a)0　　b)−2　　c)2　　d)0

63.
```
procedure generate(n: 非负整数)
if n 是奇数 then
begin
   S := S(n−1); T := T(n−1)
end
else if n=0 then
begin
   S := ∅; T := {λ}
end
else
begin
   T₁ := T(n−2); S₁ := S(n−2)
   T := T₁ ∪ {(x) | x ∈ T₁ ∪ S₁ 并且 l(x)=n−2}
   S := S₁ ∪ {xy | x ∈ T₁ 并且 y ∈ T₁ ∪ S₁ 并且 l(xy)=n}
end{T∪S 是长度至多为 n 的平衡串的集合}
```

65. 若起初 $x\leqslant y$，则不执行 $x:=y$，所以 $x\leqslant y$ 是真的终结断言。若起初 $x>y$，则执行$x:=y$，所以 $x\leqslant y$ 还是真的终结断言。

67.
```
procedure zerocount(a₁, a₂, …, aₙ: 整数表)
if n=1 then
```

```
    if a_1 = 0 then zerocount(a_1, a_2, ..., a_n) := 1
    else zerocount(a_1, a_2, ..., a_n) := 0
  else
  if a_n = 0 then zerocount(a_1, a_2, ..., a_n) := zerocount(a_1, a_2, ..., a_{n-1}) + 1
  else zerocount(a_1, a_2, ..., a_n) := zerocount(a_1, a_2, ...a_{n-1})
```

69. 将证明 $a(n)$是自然数并且 $a(n)\leqslant n$。对基础情形 $n=0$ 来说，这个命题为真，因为 $a(0)=0$。现在假设 $a(n-1)$是自然数并且 $a(n-1)\leqslant n-1$。则 $a(a(n-1))$是 a 应用到小于或等于 $n-1$ 的自然数上。因此 $a(a(n-1))$也是小于或等于 $n-1$ 的自然数。因此 $n-a(a(n-1))$是 n 减去某个小于或等于 $n-1$ 的自然数，结果是小于或等于 n 的自然数。

71. 根据练习 70，有 $a(n)=\lfloor(n+1)\mu\rfloor$和$a(n-1)=\lfloor n\mu\rfloor$。因为 $\mu<1$，所以这两个值相等或相差 1。首先假设$\mu n-\lfloor\mu n\rfloor<1-\mu$。这等价于 $\mu(n+1)<1+\lfloor\mu n\rfloor$。若这个假设为真，则$\lfloor\mu(n+1)\rfloor=\lfloor\mu n\rfloor$。在另一方面，若 $\mu n-\lfloor\mu n\rfloor\geqslant 1-\mu$，则 $\mu(n+1)\geqslant 1+\lfloor\mu n\rfloor$，所以$\lfloor\mu(n+1)\rfloor=\lfloor\mu n\rfloor+1$，即为所求。

73. $f(0)=1$，$m(0)=0$；$f(1)=1$，$m(1)=0$；$f(2)=2$，$m(2)=1$；$f(3)=2$，$m(3)=2$；$f(4)=3$，$m(4)=2$；$f(5)=3$，$m(5)=3$；$f(6)=4$，$m(6)=4$；$f(7)=5$，$m(7)=4$；$f(8)=5$，$m(8)=5$；$f(9)=6$，$m(9)=6$。

75. n 的最后一次出现在这样的位置上：1，2，…，n 的总数求和就是那个位置数。但是因为 a_k 是 k 的出现次数，所以这个数就是$\sum_{j=1}^{n}a_k$，即为所求。因为 $f(n)$ 是这个序列的前 n 项之和，所以 $f(f(n))$ 是这个序列的前 $f(n)$ 项之和。但是因为 $f(n)$ 是值为 n 的最后一项，所以这意味着这个和是这个序列里所有值不超过 n 的项之和。因为这个序列里值为 k 的项有 a_k 个，所以这个和是$\sum_{k=1}^{n}k\cdot a_k$，即为所求。

77. 假设 $f(n)$ 和 $g(n)$ 都是属于 L 的而且都是逐渐正的。则 $\ln f(n)$ 和 $\ln g(n)$ 都是属于$\mathcal{L}$的，所以根据练习 76，$\ln f(n)+\ln g(n)=\ln(f(n)g(n))\in\mathcal{L}$。因此，$e^{\ln(f(n)g(n))}=f(n)g(n)\in\mathcal{L}$。同理，$\ln f(n)-\ln g(n)=\ln(f(n)/g(n))\in\mathcal{L}$，所以 $e^{\ln(f(n)g(n))}=f(n)/g(n)\in\mathcal{L}$。
现在假设 $f(n)$和 $g(n)$都是属于$\mathcal{L}$的而且都不恒等于 0。若 $f(n)$和 $g(n)$都是逐渐负的，则 $-f(n)$和$-g(n)$都是逐渐正的而且都是属于$\mathcal{L}$的，所以 $f(n)g(n)=(-f(n))(-g(n))$和 $f(n)/g(n)=(-f(n))/(-g(n))$都是属于$\mathcal{L}$的。若这两个函数中的一个（比方说 $f(n)$）是逐渐正的，而另一个是逐渐负的，则 $f(n)g(n)=-f(n)(-g(n))\in\mathcal{L}$并且对商来说是类似的。

79. 若 $f(n)\in\mathcal{L}$并且是逐渐正的，则 $\ln f(n)\in\mathcal{L}$，所以根据练习 77 的结果，$\frac{1}{2}\ln f(n)=\ln(\sqrt{f(n)})\in\mathcal{L}$。因此$\sqrt{f(n)}=e^{\ln\sqrt{f(n)}}\in\mathcal{L}$。

第 4 章

4.1 节

1. a) 5850　　b) 343
3. a) 4^{10}　　b) 5^{10}
5. 42
7. 26^3
9. 676
11. 2^8
13. $n+1$(包含空串)
15. 475 255(包含空串)
17. 1 321 368 961

19. a)128 b)450 c)9 d)675 e)450 f)450 g)225 h)75
21. a)990 b)500 c)27
23. 3^{50}
25. 52 457 600
27. 20 077 200
29. a)37 822 859 361 b)8 204 716 800 c)40 159 050 880 d)12 113 640 000
e)171 004 205 215 f)72 043 541 640 g)6 230 721 635 h)223 149 655
31. a)0 b)120 c)720 d)2520
33. a)若 $n=1$，为 2；若 $n=2$，为 2；若 $n\geqslant 3$，为 0
b)对于 $n>1$，为 2^{n-2}；若 $n=1$，为 1 c)$2(n-1)$
35. $(n+1)^m$
37. 若 n 是偶数，为 $2^{n/2}$；若 n 是奇数，为 $2^{(n+1)/2}$
39. a)240 b)480 c)360
41. 352
43. 147
45. 33
47. 7 104 000 000 000
49. 18
51. 17
53. 22
55. 设 $P(m)$ 是关于 m 个任务的求和法则。对于基础情形取 $m=2$，这就是两个任务的求和法则。现在假设 $P(m)$ 为真，考虑 $m+1$ 个任务，T_1，T_2，…，T_{m+1}，它们可以分别用 n_1，n_2，…，n_{m+1} 种方式完成，并且没有 2 个任务是同时做的。为了完成这些任务中的 1 个任务，我们可以完成前 m 个任务中的 1 个，或者完成任务 T_{m+1}。根据两个任务的求和法则，完成任务的方式数是完成前 m 个任务之一的方式数与 n_{m+1} 之和。由归纳假设，这就是 $n_1+n_2+\cdots+n_m+n_{m+1}$，这正是所求的。
57. $n(n-3)/2$

4.2 节

1. 因为有 6 门课，但是只有 5 个工作日，鸽巢原理证明至少有 2 门课排在同一天。
3. a)3 b)14
5. 当一个整数被 4 除时只有 4 种可能的余数，给定 5 个整数，由鸽巢原理知至少有 2 个整数有同样的余数。
7. 设 a，$a+1$，…，$a+n-1$ 是序列中的整数。因为 $0<(a+j)-(a+k)<n$，其中 $0\leqslant k<j\leqslant n-1$，所以整数 $(a+i) \bmod n$ 是不等的，$i=0$，1，2，…，$n-1$。由于 $(a+i) \bmod n$ 有 n 个可能的值，且在这个集合中有 n 个不同的整数，每个值恰好取 1 次。所以序列中恰好有一个整数被 n 整除。
9. 4951
11. 点 (a, b, c) 和点 (d, e, f) 的连线的中点是 $((a+d)/2, (b+e)/2, (c+f)/2)$。它有整数坐标，当且仅当 a 和 d 的奇偶性相同，b 和 e 的奇偶性相同，且 c 和 f 的奇偶性相同。因为奇偶性的 3 元组有 8 种可能[例如(偶，奇，偶)]，根据鸽巢原理 9 个点中至少有 2 个点有相同的奇偶性 3 元组。这 2 点连线中点有整数坐标。
13. a)将前 8 个正整数 2 个一组分成 4 组使得每组整数之和等于 9：{1，8}，{2，7}，{3，6}和{4，5}。如果从前 8 个整数中选 5 个整数，由鸽巢原理至少 2 个取自同一组。这 2 个整数之和等于 9，正如所求。
b)不正确。例如取{1，2，3，4}。
15. 4
17. 21 251

19 a)如果在这个班里一年级学生少于9个，二年级学生少于9个，三年级学生也少于9个，每个年级的学生都不超过8个，那么学生总数至多24个，这与班里有25个学生矛盾。

b)如果一年级学生少于3个，二年级学生少于19个，三年级学生少于5个，那么至多2个一年级学生，至多18个二年级学生，至多4个三年级学生，学生总数至多是24，与班上有25个学生矛盾。

21. 4，3，2，1，8，7，6，5，12，11，10，9，16，15，14，13

23.
```
procedure long(a_1, …, a_n: 正整数)
{首先找最长的递增子序列}
max := 0; set := 00…00{n 位}
for i := 1 to 2^n
begin
  last := 0; count := 0; OK := true
  for j := 1 to n
  begin
    if set(j) = 1 then
    begin
      if a_j > last then last := a_j
      count := count + 1
    end
    else OK := false
  end
  if count > max then
  begin
    max := count
    best := set
  end
  set := set + 1(二进制加)
end{max 是长度并且 best 给出这个序列}
{对于递减子序列重复进行，唯一不同的是用 a_j < last 代替 a_j > last 且用 last: = ∞代替 last: = 0}
```

25. 根据对称性我们只需证明第一个语句。设 A 是其中的一个人。A 在其他9个人中或至少有4个朋友，或至少有6个敌人(因为 $3+5<9$)。假设是第一种情况，B，C，D 和 E 是 A 的朋友。如果这些人中任有2个人是朋友，我们已经找到3个人彼此都是朋友。否则 $\{B, C, D, E\}$ 是彼此都是敌人的4人集合。在第二种情况下，设 $\{B, C, D, E, F, G\}$ 是 A 的敌人的集合。由例11，在 B，C，D，E，F 和 G 中有3个人彼此都是朋友，或有3个人彼此都是敌人，这样，他们和 A 构成4人的彼此都是敌人的集合。

27. 我们需要证明两件事：如果我们有一组 n 个人，那么在他们中间一定可找到一对朋友，或者这些人的一个 n 元子集使得彼此都是敌人；并且存在一组 $n-1$ 个人，对于他们这是不可能的。对第一个陈述，若存在一对朋友，那么这个条件被满足，若不是，那么每一对人都是敌人，因此第二个条件被满足。对于第二个陈述，如果我们有 $n-1$ 个人，他们彼此都是敌人，那么既没有一对朋友，也没有他们的一个 n 元子集使得所有人彼此都是敌人。

29. 对于姓名的3个缩写字母和生日有6 432 816种可能性，因此，由鸽巢原理存在至少 $\lceil 34\,000\,000/6\,432\,816 \rceil = 6$ 个人，他们姓名的3个缩写字母和生日都相同。

31. 18

33. 因为有6台计算机，连接到同一台计算机的其他机器数是0～5之间的整数，包含0和5在内。但是0和5不能同时出现。为此只需注意到下面的事实：如果某台计算机不与其他机器连接，那么没有1台计算机连接到所有其他的5台机器；如果某台计算机连接到所有其他的5台机器，那么就没有不与其他机器相连的计算机。因为1台计算机连接的机器数至多有5种可能，于是，由鸽巢原理在6台计算

机中至少有 2 台连接的其他机器数相等。

35. 将计算机标号为 C_1 到 C_{100}，且将打印机标号为 P_1 到 P_{20}。如果对于 $k=1$，2，…，20，我们把 C_k 连接到 P_k，并且把 C_{21} 到 C_{100} 的每台计算机连接到每台打印机，那么我们总共用了 $20+80\cdot 20=1620$ 条缆线。很显然这是足够的，因为如果计算机 C_1 到 C_{20} 需要打印机，那么它们可以使用同样标号的打印机；如果任何带有更高标号的计算机需要打印机，那么它们可以使用那些没被用到的打印机，因为它们被连接到所有的打印机。现在，我们必须证明 1619 条缆线是不够用的。因为有 1619 条缆线和 20 台打印机，每台打印机平均连接的计算机数是 1619/20，它小于 81。因此某台打印机一定连接少于 81 台的计算机。这意味着它连接到 80 台或者更少的计算机，因此存在 20 台计算机没有与它连接。如果所有这 20 台计算机同时需要一台打印机，那么它们的运气就不好了，因为它们连接到至多 19 台打印机。

37. 设 a_i 是到第 i 小时为止所完成的比赛数，那么 $1\leqslant a_1<a_2<\cdots<a_{75}\leqslant 125$。进而有 $25\leqslant a_1+24<a_2+24<\cdots<a_{75}+24\leqslant 149$。有 150 个数 a_1，a_2，…，a_{75}，a_1+24，a_2+24，…，$a_{75}+24$。由鸽巢原理至少有 2 个数相等。由于所有的 ai 是不等的，所有的 a_i+24 也是不等的，因此对于某个 $i>j$ 有 $a_i=a_j+24$。于是在从第 $j+1$ 到第 i 小时之间恰有 24 场比赛。

39. 使用广义鸽巢原理，对于 $s\in S$，把 $|S|$ 个物体 $f(s)$ 放到 $|T|$ 个盒子里，T 中的每个元素一个盒子。

41. 设 d_j 是 $jx-N(jx)$，其中 $N(jx)$ 是距 jx 最近的整数，$1\leqslant j\leqslant n$。每个 d_j 是在 $-1/2$ 和 $1/2$ 之间的无理数。假设 n 是偶数，n 是奇数的情况更凌乱一些。考虑 n 个区间 $\{x\mid j/n<x<(j+1)/n\}$，$\{x\mid -(j+1)/n<x<-j/n\}$，$j=0$，1，…，$(n/2)-1$。如果对于某个 j，d_j 属于区间 $\{x\mid 0<x<1/n\}$ 或区间 $\{x\mid -1/n<x<0\}$，命题得证。如果不是，则因为有 $n-2$ 个区间和 n 个数 d_j，由鸽巢原理存在一个区间 $\{x\mid (k-1)/n<x<k/n\}$ 包含 d_r 和 d_s，其中 $r<s$。只要证明 $(s-r)x$ 距最近的整数在 $1/n$ 之内这个证明就完成了。

43. a)假设对所有的 k，$ik\leqslant n$，那么由广义鸽巢原理，数 i_1，i_2，…，i_{n^2+1} 中至少有 $\lceil (n^2+1)/n\rceil=n+1$ 个数相等。

    b)如果 $a_{k_j}<a_{k_{j+1}}$，那么 a_{k_j} 和以 $a_{k_{j+1}}$ 开始长为 $i_{k_{j+1}}$ 的递增子序列组成的子序列与 $i_{k_j}=i_{k_{j+1}}$ 的事实矛盾。于是，$a_{k_j}>a_{k_{j+1}}$。

    c)如果没有长度大于 n 的递增子序列，那么使用 a)和 b)，从而得到长度为 $n+1$ 的递减子序列 $a_{k_1}>a_{k_2}>\cdots>a_{k_{n+1}}$。

4.3 节

1. abc，acb，bac，bca，cab，cba
3. 720
5. a)120　b)720　c)8　d)6720　e)40 320　f)3 628 800
7. 15 120
9. 1320
11. a)210　b)386　c)848　d)252
13. $2(n!)^2$
15. 65 780
17. $2^{100}-5051$
19. a)1024　b)45　c)176　d)252
21. a)120　b)24　c)120　d)24　e)6　f)0
23. 609 638 400
25. a)94 109 400　b)941 094　c)3 764 376　d)90 345 024　e)114 072　f)2328　g)24　h)79 727 040　i)3 764 376　j)109 440
27. a)12 650　b)303 600
29. a)37 927　b)18 915
31. a)122 523 030　b)72 930 375　c)223 149 655　d)100 626 625

33. 54 600

35. 45

37. 912

39. 11 232 000

41. 13

43. 873

4.4 节

1. $x^4+4x^3y+6x^2y^2+4xy^3+y^4$

3. $x^6+6x^5y+15x^4y^2+20x^3y^3+15x^2y^4+6xy^5+y^6$

5. 101

7. $-2^{10}\binom{19}{9}=-94\,595\,072$

9. $-2^{101}3^{99}\binom{200}{99}$

11. 若 $k\equiv 2(\bmod 3)$ 且 $-100\leqslant k\leqslant 200$，为 $(-1)^{(200-k)/3}\binom{100}{(200-k)/3}$，否则为 0

13. 1 9 36 84 126 126 84 36 9 1

15. k 从 0 到 n 的所有正数 $\binom{n}{k}$ 的和是 2^n，因此它们中的每一项都不大于这个和。

17. $\binom{n}{k}=\frac{n(n-1)(n-2)\cdots(n-k+1)}{k(k-1)(k-2)\cdots 2}\leqslant\frac{n\cdot n\cdot\cdots\cdot n}{2\cdot 2\cdot\cdots\cdot 2}=n^k/2^{k-1}$

19. $\binom{n}{k-1}+\binom{n}{k}=\frac{n!}{(k-1)!\ (n-k+1)!}+\frac{n!}{k!\ (n-k)!}=\frac{n!}{k!\ (n-k+1)!}[k+(n-k+1)]=\frac{(n+1)!}{k!\ (n+1-k)!}=\binom{n+1}{k}$

21. a)我们证明每边计数了从一个 n 元素集合中选一个 k 个元素的子集和这个子集的一个特殊元素的方法数。对于左边，首先选择这个 k 子集[这可以用 $\binom{n}{k}$ 种方法做到]，然后在这个子集的 k 个元素中选择一个元素作为特殊的元素(这可以用 k 种方法做到)。对于右边，首先从整个 n 元素集合中选择一个特殊的元素[这可以用 n 种方法做到]，然后从剩下的 $n-1$ 个元素中选择这个子集的剩下的 $k-1$ 个元素(这可以用 $\binom{n-1}{k-1}$ 种方法做到)。

    b) $k\binom{n}{k}=k\cdot\frac{n!}{k!\ (n-k)!}=\frac{n\cdot(n-1)!}{(k-1)!\ (n-k)!}=n\binom{n-1}{k-1}$

23. $\binom{n+1}{k}=\frac{(n+1)!}{k!\ (n+1-k)!}=\frac{(n+1)}{k}\frac{n!}{(k-1)!\ (n-(k-1))!}=(n+1)\binom{n}{k-1}/k$

    这个恒等式与 $\binom{n}{0}=1$ 一起给出了一个递归定义。

25. $\binom{2n}{n+1}+\binom{2n}{n}=\binom{2n+1}{n+1}=\frac{1}{2}\left(\binom{2n+1}{n+1}+\binom{2n+1}{n+1}\right)=\frac{1}{2}\left(\binom{2n+1}{n+1}+\binom{2n+1}{n}\right)=\frac{1}{2}\binom{2n+2}{n+1}$

27. a) $\binom{n+r+1}{r}$ 通过选择 0 的位置计数了含 r 个 0 和 $n+1$ 个 1 的序列个数。另一方面，假设第 $j+1$ 项是等于 1 的最后一项，因此 $n\leqslant j\leqslant n+r$。一旦确定了最后一个 1 的位置，我们就可以确定在最后一个 1

前边的 j 个位置中有哪些位置放 0，在这个范围内有 n 个 1 和 $j-n$ 个 0。根据求和法则有 $\sum_{j=n}^{n+r}\binom{j}{j-n}=\sum_{k=0}^{r}\binom{n+k}{k}$ 种方式做这件事。

b)设 $P(r)$ 是需证明的命题。基础步骤是等式 $\binom{n}{0}=\binom{n+1}{0}$，它就是 $1=1$。假设 $P(r)$ 为真，那么 $\sum_{k=0}^{r+1}\binom{n+k}{k}=\sum_{k=0}^{r}\binom{n+k}{k}+\binom{n+r+1}{r+1}=\binom{n+r+1}{r}+\binom{n+r+1}{r+1}=\binom{n+r+2}{r+1}$，由归纳假设和帕斯卡等式得证。

29. 首先以 n 种不同的方式选择领导，然后以 2^{n-1} 种方式选择委员会的其他成员，于是有 $n2^{n-1}$ 种方式选择委员会及其领导。另一方面，选择一个 k 人委员会的方式数是 $\binom{n}{k}$。一旦我们选择了 k 个人的委员会，存在 k 种方式选择它的领导，于是有 $\sum_{k=1}^{n}k\binom{n}{k}$ 种方式选择这个委员会及其领导。从而有 $\sum_{k=1}^{n}k\binom{n}{k}=n2^{n-1}$。

31. 设这个集合有 n 个元素。由推论 2 有 $\binom{n}{0}-\binom{n}{1}+\binom{n}{2}-\cdots+(-1)^n\binom{n}{n}=0$。从而得到 $\binom{n}{0}+\binom{n}{2}+\binom{n}{4}+\cdots=\binom{n}{1}+\binom{n}{3}+\binom{n}{5}+\cdots$。左边给出了具有偶数个元素的子集个数，且右边给出了具有奇数个元素的子集个数。

33. a)一条所求的路径由 m 次向右的移动和 n 次向上的移动构成。每一条这种路径可以由一个 m 个 0 和 n 个 1 的 $m+n$ 位二进制串表示，其中 0 表示一次向右的移动，1 表示一次向上的移动。

b)恰包含 n 个 1 的 $m+n$ 位二进制串个数等于 $\binom{m+n}{n}=\binom{m+n}{m}$，因为可以通过指定 n 个 1 的位置或指定 m 个 0 的位置确定这样的串。

35. 练习 33 所描述的 n 步长的路径条数等于 2^n，就是 n 位二进制串的个数。另一方面，练习 33 中所描述的 n 步长的路径一定在坐标之和等于 n 的某个点结束，比如说$(n-k, k)$点，其中 k 在 0 和 n 之间，包含 0 和 n 在内。由练习 33，这种在$(n-k, k)$点结束的路径数等于 $\binom{n-k+k}{k}=\binom{n}{k}$。于是 $\sum_{k=0}^{n}\binom{n}{k}=2^n$。

37. 练习 33 所描述的从$(0, 0)$到$(n+1, r)$的路径数等于 $\binom{n+r+1}{r}$。但是这种路径由向上走 j 步开始，其中 j 满足 $0\leqslant j\leqslant r$。由向上走 j 步开始的这种路径数等于练习 33 所描述的从$(1, j)$到$(n+1, r)$的路径数，而这又与从$(0, 0)$到$(n, r-j)$的路径数一样，根据练习 33 它等于 $\binom{n+r-j}{r-j}$。因为 $\sum_{j=0}^{r}\binom{n+r-j}{r-j}=\sum_{k=0}^{r}\binom{n+k}{k}$，从而 $\sum_{k=1}^{r}\binom{n+k}{k}=\binom{n+r+1}{r}$。

39. a) $\binom{n+1}{2}$ b) $\binom{n+2}{3}$ c) $\binom{2n-2}{n-1}$ d) $\binom{n-1}{\lfloor (n-1)/2\rfloor}$

e)在帕斯卡三角的第 n 行中的最大的奇数项 f) $\binom{3n-3}{n-1}$

4.5 节

1. 243
3. 26^6
5. 125
7. 35

9. a)1716　b)50 388　c)2 629 575　d)330　e)9724
11. 9
13. 4 504 501
15. a)10 626　b)1365　c)11 649　d)106
17. 2520
19. 302 702 400
21. 3003
23. 7 484 400
25. 30 492
27. $C(59, 50)$
29. 35
31. 83 160
33. 63
35. 19 635
37. 210
39. 27 720
41. $52!/(7!^5 17!)$
43. $24 \cdot 13^4/(52 \cdot 51 \cdot 50 \cdot 49)$
45. a)$C(k+n-1, n)$　b)$(k+n-1)!/(k-1)!$
47. 对于第 1 个盒子有 $C(n, n_1)$种方式选择 n_1 个物体。一旦选了这些物体以后，对于第 2 个盒子有 $C(n-n_1, n_2)$种方式选择物体。类似地，对于第 3 个盒子有 $C(n-n_1-n_2, n_3)$种方式选择物体。按照这种方式继续下去直到有 $C(n-n_1-n_2-\cdots-n_{k-1}, n_k)=C(n_k, n_k)=1$ 种方式选择最后一个盒子的物体(由于 $n_1+n_2+\cdots+n_k=n$)。由乘积法则，构成全部分配的方式数是 $C(n, n_1)C(n-n_1, n_2)C(n-n_1-n_2, n_3)\cdots C(n-n_1-n_2-\cdots-n_{k-1}, n_k)$，它等于 $n!/(n_1!\,n_2!\cdots n_k!)$，正如直接化简所显示的。
49. a)由于 $x_1 \leqslant x_2 \leqslant \cdots \leqslant x_r$，从而 $x_1+0<x_2+1<\cdots<x_r+r-1$。这个不等式是严格的，因为只要 $x_j \leqslant x_{j+1}$ 就有 $x_j+j-1<x_{j+1}+j$。因为 $1 \leqslant x_j \leqslant n+r-1$，这个序列是由从 T 中选择 r 个不同的元素构造出来的。
   b)假设 $1 \leqslant x_1<x_2<\cdots<x_r \leqslant n+r-1$。令 $y_k=x_k-(k-1)$。那么不难看出对于 $k=1, 2, \cdots, r-1$ 有 $y_k \leqslant y_{k+1}$，且对于 $k=1, 2, \cdots, r$ 有 $1 \leqslant y_k \leqslant n$。从而$\{y_1, y_2, \cdots, y_k\}$是 S 的允许重复的 r-组合。
   c)从 a)和 b)得出，在 S 的允许重复的 r-组合与集合 T 的具有 $n+r-1$ 个元素的 r-组合之间存在一一对应。可以断言存在着 $C(n+r-1, r)$个 S 的允许重复的 r-组合。
51. 5
53. 在展开式中的项形如 $x_1^{n_1} x_2^{n_2} \cdots x_m^{n_m}$，其中 $n_1+n_2+\cdots+n_m=n$。产生这样的项是从 n_1 个因式中选择了 x_1，从 n_2 个因式中选择了 x_2，从 n_m 个因式中选择了 x_m，这可以用 $C(n; n_1, n_2, \cdots, n_m)$种方式做到，因为一种选择就是 n_1 个标签“1”，n_2 个标签“2”……和 n_m 个标签“m”的排列。
55. 2520

4.6 节

1. a)2134　b)54132　c)12534　d)45312　e)6714253　f)31542678
3. 14532，15432，21345，23451，23514，31452，31542，43521，45213，45321
5. 1234，1243，1324，1342，1423，1432，2134，2143，2314，2341，2413，2431，3124，3142，3214，3241，3412，3421，4123，4132，4213，4231，4312，4321
7. {1，2，3}，{1，2，4}，{1，2，5}，{1，3，4}，{1，3，5}，{1，4，5}，{2，3，4}，{2，3，5}，{2，4，5}，{3，4，5}
9. 表示下一个最大的 r-组合的二进制串一定与表示原来的 r-组合的二进制串在第 i 位不同，因为第 $i+1$，…，r

位被最大可能的数占据。此外，如果我们想要一个比原来的组合大的组合，a_i+1 是可以放在第 i 位的最小可能的数。那么 a_i+2，…，$a_i+r-i+1$ 是从第 $i+1$ 到 r 位所允许的最小的数。于是我们生成了下一个 r-组合。

11. 123，132，213，231，312，321，124，142，214，241，412，421，125，152，215，251，512，521，134，143，314，341，413，431，135，153，315，351，513，531，145，154，415，451，514，541，234，243，324，342，423，432，235，253，325，352，523，532，245，254，425，452，524，542，345，354，435，453，534，543

13. 我们将通过证明它有一个逆来证明这是一个双射。给定一个小于 $n!$ 的正整数，设 a_1，a_2，…，a_{n-1} 是它的康托数字。把 n 放在位置 $n-a_{n-1}$，因此很清楚 a_{n-1} 是在排列中跟在 n 后面且小于 n 的整数个数。然后把 $n-1$ 放在空位置 $(n-1)-a_{n-2}$，这里我们已经标号了空位置 1，2，…，$n-1$(除去 n 已经在的位置)。继续下去直到 1 放在留下的唯一的空位置。因为我们已经构造了一个逆，这个对应是一个双射。

15.
```
procedure Cantor permutation(n, i: 整数满足 n≥1 和 0≤i<n!)
x := n
for j := 1to n
  p_j := 0
for k := 1to n-1
begin
  c := ⌊x/(n-k)!⌋; x := x-c(n-k)!; h := n
  while p_h ≠ 0
    h := h-1
  for j := 1to c
  begin
    h := h-1
    while p_h ≠ 0
      h := h-1
   end
   p_h := n-k+1
  end
  h := 1
  while p_h ≠ 0
    h := h+1
    p_h := 1
      {p_1 p_2 … p_n 是对应于 i 的排列}
```

补充练习

1. a)151 200　b)1 000 000　c)210　d)5005

3. 3^{100}

5. 24 600

7. a)4060　b)2688　c)25 009 600

9. a)192　b)301　c)300　d)300

11. 639

13. 最大可能的和是 240，且最小可能的和是 15。所以可能的和的个数是 226。因为一个 10 元素集合存在 252 个 5 元子集，根据鸽巢原理至少 2 个有同样的和。

15. a)50　b)50　c)14　d)5

17. 设 a_1，a_2，…，a_m 是整数，且令 $d_i=\sum_{j=1}^{i}a_j$ 。如果对于某个 i，$d_i\equiv 0(\bmod n)$，命题得证。否则 $d_1 \bmod m$，$d_2 \bmod m,\cdots,d_m \bmod m$ 是 m 个在 $\{1,2,\cdots,m-1\}$ 中的整数。由鸽巢原理对于某个 $1\leqslant k<l\leqslant m$ 有 $d_k=d_l$，于是 $\sum_{j=k+1}^{l}a_j=d_l-d_k\equiv 0(\bmod m)$ 。

19. 有理数 a/b 的十进制展开式可以通过 b 除 a 得到，其中 a 写成含有小数点的十进制小数，且后面跟随着无数个 0。基本步骤是找出商的下一个十进制数字，即 $\lfloor r/b\rfloor$，其中 r 是余数和从被除数移下来的下一位。当前的余数由前面的余数减去商的前面数字的 b 倍得到。最后被除数没有数字时移下 0。此外，由于只存在 b 个可能的余数。于是由鸽巢原理，在某一步将得到与前面某一步相同的结果。从这一步向后计算一定遵循着相同的模式。特别是商将被重复。

21. a)125 970　b)20　c)141 120 525　d)141 120 505　e)177 100　f)141 078 021

23. a)10　b)8　c)7

25. 3^n

27. $C(n+2, r+1)=C(n+1, r+1)+C(n+1, r)=2C(n+1, r+1)-C(n+1, r+1)+C(n+1, r)=2C(n+1, r+1)-(C(n, r+1)+C(n, r))+(C(n, r)+C(n, r-1))=2C(n+1, r+1)-C(n, r+1)+C(n, r-1)$

29. 把 $x=1$ 和 $y=3$ 代入二项式定理。

31. $C(n+1, 5)$

33. 3 491 888 400

35. 5^{24}

37. a)45　b)57　c)12

39. a)386　b)56　c)512

41. 如果 $n<m$ 为 0；如果 $n\geqslant m$ 为 $C(n-1, n-m)$。

43. **procedure** *next permutation*(n：正整数，a_1，a_2，…，a_r：不超过 n 的正整数具有 $a_1a_2\cdots a_r\neq nn\cdots n$)

```
i := r
while ai = n
begin
    a_i := 1
    i := i - 1
end
a_i := a_i + 1
{a1 a2 … ar 是按照字典次序的下一个排列}
```

45. 我们必须证明，如果在一个晚会上存在 $R(m, n-1)+R(m-1, n)$ 个人，那么一定至少有 m 个人彼此是朋友，或者 n 个人彼此是敌人。考虑一个人，让我们称他为杰瑞。那么在晚会上有 $R(m-1, n)+R(m, n-1)-1$ 个其他的人，并且根据鸽巢原理，在这些人中一定是或者至少有 $R(m-1, n)$ 个杰瑞的朋友，或者 $R(m, n-1)$ 个杰瑞的敌人。首先假定存在 $R(m-1, n)$ 个杰瑞的朋友，根据 R 的定义，在这些人中，我们保证发现 $m-1$ 个人彼此都是朋友，或者 n 个人彼此都是敌人。在前一种情况下，这 $m-1$ 个彼此的朋友与杰瑞一起就是 m 个彼此是朋友的人的集合；在后一种情况下，我们有了所要求的 n 个彼此是敌人的人的集合。其他的情况类似：假设存在 $R(m, n-1)$ 个杰瑞的敌人，我们保证在他们中间发现或者 m 个人彼此是朋友，或者 $n-1$ 个人彼此是敌人。在前一种情况下，我们有所要求的 m 个人彼此都是朋友；在后一种情况下，这 $n-1$ 个彼此的敌人与杰瑞一起是 n 个彼此是敌人的人的集合。

第 5 章

5.1 节

1. 1/13

3. 1/2

5. 1/2

7. 1/64

9. 47/52

11. 1/C(52，5)

13. 1−(C(48，5)/C(52，5))

15. C(13，2)C(4，2)C(4，2)C(44，1)/C(52，5)

17. 10 240/C(52，5)

19. 1 302 540/C(52，5)

21. 1/64

23. 8/25

25. a)1/C(50，6)=1/15 890 700　b)1/C(52，6)=1/20 358 520
c)1/C(56，6)=1/32 468 436　d)1/C(60，6)=1/50 063 860

27. a)139 128/319 865　b)212 667/511 313　c)151 340/386 529　d)163 647/446 276

29. 1/C(100，8)

31. 3/100

33. a)1/788 0400　b)1/8 000 000

35. a)9/19　b)81/361　c)1/19
d)1 889 568/2 476 099　e)48/361

37. 掷 3 个骰子。

39. 这个竞争者选的门是在不知道大奖在何处时随机选的，但是由主持人选的门不是随机选的，因为他总是避免打开有大奖的门。这就使得基于对称性的论据是无效的。

41. a)671/1296　b)$1-35^{24}/36^{24}$；否　c)是

5.2 节

1. $p(T)=1/4$，$p(H)=3/4$

3. $p(1)=p(3)=p(5)=p(6)=1/16$；$p(2)=p(4)=3/8$

5. 9/49

7. a)1/2　b)1/2　c)1/3　d)1/4　e)1/4

9. a)1/26!　b)1/26　c)1/2　d)1/26　e)1/650　f)1/15 600

11. 很显然，$p(E\cup F)\geqslant p(E)=0.7$。$p(E\cup F)\leqslant 1$。如果使用 5.1 节的定理 2，可以将它重写为 $p(E)+p(F)-P(E\cap F)\leqslant 1$，或者 $0.7+0.5-p(E\cap F)\leqslant 1$。求解 $p(E\cap F)$ 得到 $p(E\cap F)\geqslant 0.2$。

13. 由于 $p(E\cup F)=p(E)+p(F)-p(E\cap F)$ 且 $p(E\cup F)\leqslant 1$，从而得到 $1\geqslant p(E)+p(F)-p(E\cap F)$。由这个不等式可以断言 $p(E)+p(F)\leqslant 1+p(E\cap F)$。

15. 我们将使用数学归纳法证明这个不等式对 $n\geqslant 2$ 成立。设 $P(n)$ 是命题 $p\left(\bigcup_{j=1}^{n}E_j\right)\leqslant\sum_{j=1}^{n}p(E_j)$。
基础步骤：$P(2)$ 为真，因为 $p(E_1\cup E_2)=p(E_1)+p(E_2)-p(E_1\cap E_2)\leqslant p(E_1)+p(E_2)$。
归纳步骤：假设 $P(k)$ 为真。使用基础情况和归纳假设，从而得出 $p\left(\bigcup_{j=1}^{k+1}E_j\right)\leqslant p\left(\bigcup_{j=1}^{k}E_j\right)+p(E_{k+1})\leqslant\sum_{j=1}^{k+1}p(E_j)$，这就证明了 $P(k+1)$ 为真。根据数学归纳法命题得证。

17. 因为 $E\cup\bar{E}$ 是整个样本空间 S，事件 F 可以被分成两个不相交的事件：$F=S\cap F=(E\cup\bar{E})\cap F=(E\cap F)\cup(\bar{E}\cap F)$，这里使用了分配律。由于这两个事件是不相交的，故 $p(F)=p((E\cap F)\cup(\bar{E}\cap F))=p(E\cap F)+p(\bar{E}\cap F)$，使用事实 $p(E\cap F)=p(E)\cdot p(F)$（已知 E 和 F 是独立的）在两边替换 $p(E\cap F)$ 并且分解因式得 $p(F)(1-p(E))=p(\bar{E}\cap F)$。由于 $1-p(E)=p(\bar{E})$，这就得到 $p(\bar{E}\cap F)=p(\bar{E})\cdot p(F)$，正如所求。

19.（假设出生是独立的并且每个月出生的概率都是 1/12。）

a)1/12　　b)$1-\frac{11}{12}\cdot\frac{10}{12}\cdot\cdots\cdot\frac{13-n}{12}$　　c)5

21. 614

23. 1/4

25. 3/8

27. a)不独立　　b)不独立　　c)不独立

29. 3/16

31. a)$1/32=0.031\,25$　b)$0.49^5\approx0.028\,25$　c)0.037 950 12

33. a)5/8　　b)0.627 649　　c)0.643 1

35. a)p^n　b)$1-p^n$　c)$p^n+n\cdot p^{n-1}\cdot(1-p)$　d)$1-[p^n+n\cdot p^{n-1}\cdot(1-p)]$

37. $p(\bigcup_{i=1}^{\infty}E_i)$是在$\bigcup_{i=1}^{\infty}E_i$中的每个结果$s$的$p(s)$之和。因为$E_i$是彼此不相交的，这就是在每个$E_i$中的所有结果的概率之和，即$\sum_{i=1}^{\infty}p(E_i)$。（我们可以重新排列这些被加数并且仍旧得到同样的结果，因为这个数列是绝对收敛的。）

39. a)$\overline{E_i}=\bigcup_{j=1}^{\binom{k}{m}}F_j$，因此可以从布尔不等式(见练习 15)得到给定的不等式。

b)一个特定的不在第j个集合的游戏者赢了第j个集合的所有k个游戏者的概率是$(1/2)^k=2^{-k}$。所以，不在第j个集合的所有$m-k$个游戏者都不能以赢了第j个集合的每一个人的完美纪录而自豪的概率是$(1-2^{-k})^{m-k}$。这恰好是$p(F_j)$。

c)可以立即得到第一个不等式，因为所有的被加数都是同样的并且有$\binom{m}{k}$个被加数。如果这个概率小于 1，$\overline{E}$不发生一定是可能的，即E会发生。因此只要第二个不等式成立，就存在一个循环赛适合这个问题的条件。

d)对$k=2$有$m\geqslant21$，对$k=3$有$m\geqslant91$。

41.
```
procedure probabilistic prime (n, k)
composite := false
i := 0
while composite = false and i < k
begin
  i := i+1
  choose b uniformly at random with 1<b<n
  apply Miller's test to base b
  if n fails the test then composite = true
end
if composite = true then print ("composite")
else print("probably prime")
```

5.3 节

1. 由定理 2，$p=1/2$和$n=5$，我们看到头像的预期数目是 2.5。

3. 5/3

5. 336/49

7. 170

9. $(4n+6)/3$

11. 50，700，551/10，077，696≈5.03

13. 6

15. $p(X \geqslant j) = \sum_{k=j}^{\infty} p(X = k) = \sum_{k=j}^{\infty} (1-p)^{k-1} p = p(1-p)^{j-1} \sum_{k=0}^{\infty} (1-p)^k$

$= p(1-p)^{j-1}/(1-(1-p)) = (1-p)^{j-1}$

17. 2302

19. $(7/2) \cdot 7 \neq 329/12$

21. $p+(n-1)p(1-p)$

23. 5/2

25. a)0　　b)1

27. 1/100

29. $E(X)/a = \sum_{r} (r/a) \cdot p(X = r) \geqslant \sum_{r \geqslant a} 1 \cdot p(X = r) = p(X \geqslant a)$

31. a)10/11　b)0.9984

33. a)$n!$ 个排列每个出现的概率为 $1/n!$，所以 $E(X)$是所有这些排列的平均比较次数。

b)即使算法继续 $n-1$ 次巡回。X 将至多是 $n(n-1)/2$。从而得到关于期望的公式 $E(X) \leqslant n(n-1)/2$。

c)算法通过相邻元素的比较来继续，且如果需要的话就交换它们。原来有逆序的元素变成不是逆序的唯一方法就是对它们进行比较和交换。

d)因为对所有的 P，$X(P) \geqslant I(P)$，由期望的定义得到 $E(X) \geqslant E(I)$。

e)对一个逆序的每个实例这个和式计数一个 1。

f)由定理 3 得出。

g)在定理 2 中令 $n=1$，$I_{j,k}$的期望就是在排列中 a_k 在 a_j 前面的概率。由对称性显然是 1/2。

h)在 f)中的和式由 $C(n, 2)=n(n-1)/2$ 个项组成，每个等于 1/2，所以这个和为 $n(n-1)/4$。

i)从 a)和 b)知道 $E(X)$这个所关心的目标至多是 $n(n-1)/2$，并且由 d)和 h)知道 $E(X)$至少是 $n(n-1)/4$。两者都是 $\Theta(n^2)$。

35. 1

37. $V(X+Y)=E((X+Y)^2)-E(X+Y)^2=E(X^2+2XY+Y^2)-(E(X)+E(Y))^2$

$=E(X^2)+2E(XY)+E(Y^2)-E(X)^2-2E(X)E(Y)-E(Y)^2$

$=E(X^2)-E(X)^2+2(E(XY)-E(X)E(Y))+E(Y)^2-E(Y)^2=V(X)+2\mathrm{Cov}(X, Y)+V(Y)$

39. $((n-1)/n)^m$

41. $(n-1)^m/n^{m-1}$

补充练习

1. 1/109 668

3. a)1/C(52, 13)　　b)4/C(52, 13)　　c)2 944 656/C(52, 13)

d)35 335 872/C(52, 13)

5. a)9/2　　b)21/4

7. a)9　　b)21/2

9. a)8　　b)49/6

11. a)$n/2^{n-1}$　　b)$p(1-p)^{k-1}$，其中 $p=n/2^{n-1}$　　c)$2^{n-1}/n$

13. $\dfrac{(m-1)(n-1)+\gcd(m,n)-1}{mn-1}$

15. a)2/3　　b)2/3

17. 1/32

19. a)一个人赢 2^n 美元的概率是 $1/2^n$，因为它恰好出现在游戏者得到 $n-1$ 次头像向下，接着一次头像向上的时候。因此赢的期望值是 2^n 乘以 $1/2^n$ 的当 n 从 1 趋向无穷时的和。因为每一项是 1，这个和是无穷的。换句话说，一个人应该心甘情愿地赌任意数量的钱并且期待在长时间运行中出现头像向上。

    b)9 美元，9 美元。

21. a)当 $S=\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$，$A=\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ 且 $B=\{1, 2, 3, 4\}$ 时为 1/3；当 $S=\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$，$A=\{4, 5, 6, 7, 8, 9, 10, 11, 12\}$ 且 $B=\{1, 2, 3, 4\}$ 时为 1/12。

    b)当 $S=\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$，$A=\{4, 5, 6, 7, 8, 9, 10, 11, 12\}$ 且 $B=\{1, 2, 3, 4\}$ 时为 1；当 $S=\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$，$A=\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ 且 $B=\{1, 2, 3, 4\}$ 时为 3/4。

23. a)$p(E_1\cap E_2)=p(E_1)p(E_2)$，$p(E_1\cap E_3)=p(E_1)p(E_3)$，$p(E_2\cap E_3)=p(E_2)p(E_3)$，$p(E_1\cap E_2\cap E_3)=p(E_1)p(E_2)p(E_3)$。

    b)是　　c)是　　d)2^n-n-1

25. 事件 $E\cap F_i$，$i=1, 2, \cdots, n$，是互相排斥的并且覆盖了 E 出现的所有条件。因此，$p(E)=\sum_{i=1}^{n}p(E\cap F_i)$。由条件概率的定义可知 $p(E\cap F_i)=p(E\mid F_i)p(F_i)$ 并且 $p(E\mid F_i)=p(E\cap F_i)/p(E)$。因此

$$p(F_j\mid E)=\frac{p(E\cap F_j)}{p(E)}=\frac{p(E\mid F_j)p(F_j)}{\sum_{i=1}^{n}p(E\cap F_i)}=\frac{p(E\mid F_j)p(F_j)}{\sum_{i=1}^{n}p(E\mid F_i)p(F_i)}$$

27. $V(aX+b)=E((aX+b)^2)-E(aX+b)^2=E(a^2X^2+2abX+b^2)-(aE(X)+b)^2$
$=E(a^2X^2)+E(2abX)+E(b^2)-(a^2E(X)^2+2abE(X)+b^2)$
$=a^2E(X^2)+2abE(X)+b^2-a^2E(X)^2-2abE(X)-b^2$
$=a^2(E(X^2)-E(X)^2)=a^2V(X)$

29. 为了对样本空间的每个元素恰好计数一次，我们必须包含每个集合中的所有元素，然后去掉在交集中对元素的重复计数。因为 $C(m, 2)$ 项被减去，于是 $p(E_1\cup E_2\cup\cdots\cup E_m)=p(E_1)+p(E_2)+\cdots+p(E_m)-p(E_1\cap E_2)-p(E_1\cap E_3)-\cdots-p(E_1\cap E_m)-p(E_2\cap E_3)-p(E_2\cap E_4)-\cdots-p(E_2\cap E_m)-\cdots-p(E_{m-1}\cap E_m)=qm-(m(m-1)/2)r$，但是，$p(E_1\cup E_2\cup\cdots\cup E_m)=1$，因此有 $qm-(m(m-1)/2)r=1$。由于 $r\geqslant 0$，由这个等式知道 $qm\geqslant 1$。于是 $q\geqslant 1/m$。因为 $q\leqslant 1$，从这个等式推出 $(m(m-1)/2)r=qm-1\leqslant m-1$。从而得到 $r\leqslant 2/m$。

31. a)我们买这些卡直到每种类型的卡都得到一张。这意味着我们已经买了总共 X 张卡。另一方面，这也意味着我们买 X_0 张卡直到得到了所得的第一种类型的卡，然后多买 X_1 张卡直到得到所得的第二种类型的卡，照此下去。于是 X 是这些 X_j 的和。

    b)一旦已经得到 j 种不同类型的卡，在总数 n 种可用的类型中有 $n-j$ 种新的可用的类型。因为我们拿到每种类型是等可能的，在下一次买时成功(得到一个新类型)的概率是 $(n-j)/n$。

    c)这可以从几何分布的定义、X_j 的定义和 b)得到。

    d)从 c)得出 $E(X_j)=n/(n-j)$。于是由期望的线性性质和 a)有

$$E(X)=E(X_0)+E(X_1)+\cdots+E(X_{n-1})=\frac{n}{n}+\frac{n}{n-1}+\cdots+\frac{n}{1}=n(\frac{1}{n}+\frac{1}{n-1}+\cdots+\frac{1}{1})。$$

    e)大约 224.46。

第 6 章

6.1 节

1. a)2，12，72，432，2592　　b)2，4，16，256，65，536

c)1，2，5，11，26　　d)1，1，6，27，204　　e)1，2，0，1，3

3. a)6，17，49，143，421

b)$49=5\cdot 17-6\cdot 6$，$143=5\cdot 49-6\cdot 17$，$421=5\cdot 143-6\cdot 49$

c)$5a_{n-1}-6a_{n-2}=5(2^{n-1}+5\cdot 3^{n-1})-6(2^{n-2}+5\cdot 3^{n-2})=2^{n-2}(10-6)+3^{n-2}(75-30)$
$=2^{n-2}\cdot 4+3^{n-2}\cdot 9\cdot 5=2^n+3^n\cdot 5=a_n$

5. a)是　b)不是　c)不是　d)是　e)是　f)是　g)不是　h)不是

7. a)$a_{n-1}+2a_{n-2}+2n-9=-(n-1)+2+2(-(n-2)+2)+2n-9=-n+2=a_n$

b)$a_{n-1}+2a_{n-2}+2n-9=5(-1)^{n-1}-(n-1)+2+2(5(-1)^{n-2}-(n-2)+2)+2n-9$
$=5(-1)^{n-2}(-1+2)-n+2=a_n$

c)$a_{n-1}+2a_{n-2}+2n-9=3(-1)^{n-1}+2^{n-1}-(n-1)+2+2(3(-1)^{n-2}+2^{n-2}-(n-2)+2)+2n-9$
$=3(-1)^{n-2}(-1+2)+2^{n-2}(2+2)-n+2=a_n$

d)$a_{n-1}+2a_{n-2}+2n-9=7\cdot 2^{n-1}-(n-1)+2+2(7\cdot 2^{n-2}-(n-2)+2)+2n-9=2^{n-2}(7\cdot 2+2\cdot 7)-n+2=a_n$

9. a)$a_n=2\cdot 3^n$　b)$a_n=2n+3$　c)$a_n=1+n(n+1)/2$　d)$a_n=n^2+4n+4$
e)$a_n=1$　f)$a_n=(3^{n+1}-1)/2$　g)$a_n=5n!$　h)$a_n=2^n n!$

11. a)$a_n=3a_{n-1}$　b)5 904 900

13. a)$a_n=n+a_{n-1}$，$a_0=0$
b)$a_{12}=78$　c)$a_n=n(n+1)/2$

15. $B(k)=(1+(0.07/12))B(k-1)-100$，$B(0)=5000$

17. 设 $P(n)$是“$H_n=2^n-1$”。基础步骤：$P(1)$为真，因为 $H_1=1$。归纳步骤：假设 $H_n=2^n-1$，那么因为 $H_{n+1}=2H_n+1$，从而得到 $H_{n+1}=2(2^n-1)+1=2^{n+1}-1$。

19. a)$a_n=2a_{n-1}+a_{n-5}$，对于 $n\geqslant 5$。　b)$a_0=1$，$a_1=2$，$a_2=4$，$a_3=8$，$a_4=16$　c)1217

21. 9494

23. a)$a_n=a_{n-1}+a_{n-2}+2^{n-2}$ 对于 $n\geqslant 2$　b)$a_0=0$，$a_1=0$　c)94

25. a)$a_n=a_{n-1}+a_{n-2}+a_{n-3}$ 对于 $n\geqslant 3$　b)$a_0=1$，$a_1=2$，$a_2=4$　c)81

27. a)$a_n=a_{n-1}+a_{n-2}$ 对于 $n\geqslant 2$　b)$a_0=1$，$a_1=1$　c)34

29. a)$a_n=2a_{n-1}+2a_{n-2}$ 对于 $n\geqslant 2$　b)$a_0=1$，$a_1=3$　c)448

31. a)$a_n=2a_{n-1}+a_{n-2}$ 对于 $n\geqslant 2$　b)$a_0=1$，$a_1=3$　c)239

33. a)$a_n=2a_{n-1}$ 对于 $n\geqslant 2$　b)$a_1=3$　c)96

35. a)$a_n=a_{n-1}+a_{n-2}$ 对于 $n\geqslant 2$　b)$a_0=1$，$a_1=1$　c)89

37. a)$R_n=n+R_{n-1}$，$R_0=1$　b)$R_n=n(n+1)/2+1$

39. a)$S_n=S_{n-1}+(n^2-n+2)/2$，$S_0=1$　b)$S_n=(n^3+5n+6)/6$

41. 64

43. a)$a_n=2a_{n-1}+2a_{n-2}$　b)$a_0=1$，$a_1=3$　c)1224

45. 很显然对于 $m\geqslant 1$ 有 $S(m, 1)=1$。如果 $m\geqslant n$，那么一个从 m 元素集合到 n 元素集合的非映上函数可以由选择值域的大小和值域中的元素来确定。值域的大小是在 1 到 $n-1$ 之间的整数 k，包含 1 和 $n-1$，而这个值域可以用 $C(n, k)$种方式选取，并且选择一个到这个值域的映上函数可以有 $S(m, k)$种方式。因此存在 $\sum_{k=1}^{n-1}C(n,k)S(m,k)$ 个非映上函数。但是总共存在 n^m 个函数，所以 $S(m,n)=n^m-\sum_{k=1}^{n-1}C(n,k)S(m,k)$。

47. a)$C_5=C_0C_4+C_1C_3+C_2C_2+C_3C_1+C_4C_0=1\cdot 14+1\cdot 5+2\cdot 2+5\cdot 1+14\cdot 1=42$
b)$C(10, 5)/6=42$

49. $J(1)=1$，$J(2)=1$，$J(3)=3$，$J(4)=1$，$J(5)=3$，$J(6)=5$，$J(7)=7$，$J(8)=1$，$J(9)=3$，$J(10)=5$，$J(11)=7$，$J(12)=9$，$J(13)=11$，$J(14)=13$，$J(15)=15$，$J(16)=1$

51. 首先，假设人数是偶数，比如说是 $2n$。在转了一圈并且回到第一个人以后，由于在偶数位置的人已经被排除，恰好有 n 个人留下来并且现在在位置 i 的人就是初始在位置 $2i-1$ 的人。因此，生还者[初始在位置 $J(2n)$]现在的位置是 $J(n)$；这就是在位置 $2J(n)-1$ 的人。所以，$J(2n)=2J(n)-1$。类似地，当有奇数个人时，比如说 $2n+1$ 个人，那么在转了一圈以后排除第 1 个人，有 n 个人留下来并且现在在位置 i 的人就是以前在位置 $2i+1$ 的人。因此，生还者将是现在占据位置 $J(n)$ 的人，即初始在位置 $2J(n)+1$ 的人。所以，$J(2n+1)=2J(n)+1$。基础情形是 $J(1)=1$。

53. 73，977，3617

55. 下面的 9 次移动求解了这个难题：从柱 1 到柱 2 移动盘 1；从柱 1 到柱 3 移动盘 2；从柱 2 到柱 3 移动盘 1；从柱 1 到柱 2 移动盘 3；从柱 1 到柱 4 移动盘 4；从柱 2 到柱 4 移动盘 3；从柱 3 到柱 2 移动盘 1；从柱 3 到柱 4 移动盘 2；从柱 2 到柱 4 移动盘 1。要明白至少需要 9 次移动，首先注意，不管有多少根柱子至少 7 次移动是需要的：3 次用来拿走盘子，1 次用来移动最大的盘 4，并且还要 3 次移动来放好盘子。至少还需要 2 次另外的移动，因为要从柱 1 到柱 4 移动盘 4，其他 3 个盘子一定在柱 2 和柱 3，所以至少需要 1 次移动来放好它们和 1 次移动来拿走它们。

57. 基础情形是显然的。如果 $n>1$，算法由 3 步构成。第 1 步，由归纳假设，使用 $R(n-k)$ 次移动把最小的 $n-k$ 个盘子移到柱 2。然后使用通常的 3 个柱子的汉诺塔算法，用 2^k-1 次移动把剩下的盘子(最大的 k 个盘子)移到柱 4，移动时不使用柱 2。然后再由归纳法，用 $R(n-k)$ 次移动把最小的 $n-k$ 个盘子移到柱 4；对于这种移动，所有的柱子都可以用，因为最大的盘子在柱 4，对移动没有妨碍。这就建立了递推关系。

59. 首先注意到 $R(n)=\sum_{j=1}^{n}(R(j)-R(j-1))$(因为化简这个和并且 $R(0)=0$)。由练习 58，这就是关于 j 的值域对 $2^{k'-1}$ 求和，因此这个和是 $\sum_{i=1}^{k}i2^{i-1}$，如果 n 不是一个三角形数，那么当 $i=k$ 时最后某些值丢失，并且这就是在给定表达式中最后的项所表示的。

61. 由练习 59，$R(n)$不大于 $\sum_{i=1}^{k}i2^{i-1}$ 。可以证明这个和等于$(k+1)2^k-2^{k+1}+1$，因此它不比$(k+1)2^k$ 大。因为 $n>k(k-1)/2$，可以用二次公式证明对于所有的 $n>1$ 有 $k<1+\sqrt{2n}$。因而对于所有的 $n>2$，$R(n)$以$(1+\sqrt{2n}+1)2^{1+\sqrt{2n}}<8\sqrt{n}2^{\sqrt{2n}}$为上界。所以 $R(n)$是 $O(\sqrt{n}2^{\sqrt{2n}})$。

63. a)0　b)0　c)2　d)$2^{n-1}-2^{n-2}$

65. $a_n-2\nabla a_n+\nabla^2 a_n$
$=a_n-2(a_n-a_{n-1})+(\nabla a_n-\nabla a_{n-1})$
$=-a_n+2a_{n-1}+((a_n-a_{n-1})-(a_{n-1}-a_{n-2}))$
$=-a_n+2a_{n-1}+(a_n-2a_{n-1}+a_{n-2})$
$=a_{n-2}$

67. $a_n=a_{n-1}+a_{n-2}$
$=(a_n-\nabla a_n)+(a_n-2\nabla a_n+\nabla^2 a_n)$
$=2a_n-3\nabla a_n+\nabla^2 a_n$，
或 $a_n=3\nabla a_n-\nabla^2 a_n$

6.2 节

1. a)3 阶　b)不是　c)4 阶　d)不是　e)不是　f)2 阶　g)不是

3. a)$a_n=3\cdot 2^n$　b)$a_n=2$　c)$a_n=3\cdot 2^n-2\cdot 3^n$
d)$a_n=6\cdot 2^n-2\cdot n2^n$　e)$a_n=n(-2)^{n-1}$　f)$a_n=2^n-(-2)^n$
g)$a_n=(1/2)^{n+1}-(-1/2)^{n+1}$

5. $a_n=\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1}-\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{n+1}$

7. $(2^{n+1}+(-1)^n)/3$

9. a) $P_n=1.2P_{n-1}+0.45P_{n-2}$，$P_0=100\,000$，$P_1=120\,000$

 b) $P_n=(250\,000/3)(3/2)^n+(50\,000/3)(-3/10)^n$

11. a) 基础步骤：对于 $n=1$ 有 $1=0+1$，且对于 $n=2$ 有 $3=1+2$。归纳步骤：假设对于 $k\leqslant n$ 为真，那么 $L_{n+1}=L_n+L_{n-1}=f_{n-1}+f_{n+1}+f_{n-2}+f_n=(f_{n-1}+f_{n-2})+(f_{n+1}+f_n)=f_n+f_{n+2}$

 b) $L_n=\left(\frac{1+\sqrt{5}}{2}\right)^n+\left(\frac{1+\sqrt{5}}{2}\right)^n$

13. $a_n=8(-1)^n-3(-2)^n+4\cdot 3^n$

15. $a_n=5+3(-2)^n-3^n$

17. 设 $a_n=C(n,0)+C(n-1,1)+\cdots+C(n-k,k)$ 其中 $k=\lfloor n/2\rfloor$。首先，假设 n 是偶数，使得 $k=n/2$，且最后项是 $C(k,k)$。由帕斯卡恒等式有 $a_n=1+C(n-2,0)+C(n-2,1)+C(n-3,1)+C(n-3,2)+\cdots+C(n-k,k-2)+C(n-k,k-1)+1=1+C(n-2,1)+C(n-3,2)+\cdots+C(n-k,k-1)+C(n-2,0)+C(n-3,1)+\cdots+C(n-k,k-2)+1=a_{n-1}+a_{n-2}$，因为 $\lfloor(n-1)/2\rfloor=k-1=\lfloor(n-2)/2\rfloor$。当 n 是奇数时有类似的计算。因此，对于所有的正整数 n，$n\geqslant 2$，$\{a_n\}$ 满足递推关系 $a_n=a_{n-1}+a_{n-2}$。此外，$a_1=C(1,0)=1$ 和 $a_2=C(2,0)+C(1,1)=2$，这就是 f_2 和 f_3。从而得到对于所有的正整数 n，$a_n=f_{n+1}$。

19. $a_n=(n^2+3n+5)(-1)^n$

21. $(a_{1,0}+a_{1,1}n+a_{1,2}n^2+a_{1,3}n^3)+(a_{2,0}+a_{2,1}n+a_{2,2}n^2)(-2)^n+(a_{3,0}+a_{3,1}n)3^n+a_{4,0}(-4)^n$

23. a) $3a_{n-1}+2n=3(-2^n)+2^n=2^n(-3+1)=-2^{n+1}=a_n$

 b) $a_n=\alpha 3^n-2^{n+1}$ c) $a_n=3^{n+1}-2^{n+1}$

25. a) $A=-1$，$B=-7$ b) $a_n=\alpha 2^n-n-7$ c) $a_n=11\cdot 2^n-n-7$

27. a) $p_3n^3+p_2n^2+p_1n+p_0$ b) $n^2p_0(-2)^n$ c) $n^2(p_1n+p_0)2^n$ d) $(p_2n^2+p_1n+p_0)4^n$

 e) $n^2(p_2n^2+p_1n+p_0)(-2)^n$ f) $n^2(p_4n^4+p_3n^3+p_2n^2+p_1n+p_0)2^n$ g) p_0

29. a) $a_n=\alpha 2^n+3^{n+1}$ b) $a_n=-2\cdot 2^n+3^{n+1}$

31. $a_n=\alpha 2^n+\beta 3^n-n\cdot 2^{n+1}+3n/2+21/4$

33. $a_n=(\alpha+\beta n+n^2+n^3/6)2^n$

35. $a_n=-4\cdot 2^n-n^2/4-5n/2+1/8+(39/8)3^n$

37. $a_n=n(n+1)(n+2)/6$

39. a) 1，-1，$i-i$ b) $a_n=\frac{1}{4}-\frac{1}{4}(-1)^n+\frac{2+i}{4}i^n+\frac{2-i}{4}(-i)^n$

41. a) 使用关于 f_n 的公式可看出

$$\left|f_n-\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n\right|=\left|\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n\right|<1/\sqrt{5}<1/2$$

 这意味着 f_n 是最接近 $\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n$ 的整数。

 b) 当 n 是偶数时较小；当 n 是奇数时较大。

43. $a_n=f_{n-1}+2f_n-1$

45. a) $a_n=3a_{n-1}+4a_{n-2}$，$a_0=2$，$a_1=6$

 b) $a_n=(4^{n+1}+(-1)^n)/5$

47. a) $a_n=2a_{n-1}+(n-1)10\,000$

 b) $a_n=70\,000\ 2^{n-1}-10\,000\ n-10\,000$

49. $a_n=5n^2/12+13n/12+1$

51. 见第 11 章第 5 节参考文献[Ma93]

53. $6^n \cdot 4^{n-1}/n$

6.3 节

1. 14

3. 第一步是$(1110)_2(1010)_2=(2^4+2^2)(11)_2(10)_2+2^2((11)_2-(10)_2)((10)_2-(10)_2)+(2^2+1)(10)_2\cdot(10)_2$。这个积是$(10\ 001\ 100)_2$。

5. $C=50\ 665C+729=33\ 979$

7. a)2　　b)4　　c)7

9. a)79　　b)48 829　　c)30 517 579

11. $O(\log n)$

13. $O(n^{\log_3 2})$

15. 5

17. a)基础步骤：如果序列只有一个元素，那么在这个表中的一个人就是赢的人。递归步骤：把这个表分成两部分(前一半和后一半尽可能相等)。递归地应用算法到每一半直到至多出现两个名字为止。然后检查整个表来计算每个名字出现的数目，如果谁超过半数就决定谁赢。

   b)$O(n\log n)$

19. a)$f(n)=f(n/2)+2$

   b)$O(\log n)$

21. a)7　　b)$O(\log n)$

23. a)
```
procedure largest sum(a1, a2, …, an)
best := 0{空序列和为 0}
for i := 1 to n
begin
  sum := 0
  for j := i+1 to n
  begin
    sum := sum+aj
    if sum>best then best := sum
  end
end
{best 是在这个表中可能的最大和}
```

   b)$O(n^2)$

   c)我们把这个表划分成前一半和后一半，并且递归地对每一半应用算法找连续项的最大和。在整个序列的连续项的最大和是这两个数中的一个，或者是横跨这个表中间的连续项的序列之和。为了找到横跨这个表中间的连续项序列的最大可能的和，我们从表的中间开始向后移动来找在这个表的后一半的最大可能的和，并且向前移动来找在这个表的前一半的最大可能的和；所要的和就是这两个量之和。那么最后的答案就是这个和以及递归得到的两个答案之中的最大者。基础情况是一个项的序列的最大和就是这个数和 0 之中的较大者。

   d)11，9，14

   e)$S(n)=2S(n/2)+n$，$C(n)=2C(n/2)+n+2$，$S(1)=0$，$C(1)=1$

   f)$O(n\log n)$，比 $O(n^2)$好

25. (1，6)和(3，2)距离为 2。

27. 算法基本上与例 12 给出的算法一样。中心空隙宽度仍旧是 $2d$，但我们仅需要考虑 2 个大小 $d\times d$ 的盒子而不是大小$(d/2)\times(d/2)$的 8 个盒子。除了系数 7 被 1 替换以外，递推关系与例 12 的递推关系

一样。

29. 由于 $k=\log_b n$，从而有 $f(n)=a^k f(1)+\sum_{j=0}^{k-1} a^j c(n/b^j)^d = a^k f(1)+\sum_{j=0}^{k-1} cn^d = a^k f(1)+kcn^d = a^{\log_b n} f(1)+c(\log_b n)n^d = n^{\log_b a} f(1)+cn^d \log_b n = n^d f(1)+cn^d \log_b n$。

31. 设 $k=\log_b n$，其中 n 是 b 的幂。基础步骤：如果 $n=1$ 且 $k=0$，那么 $c_1 n^d + c_2 n^{\log_b a} = c_1 + c_2 = b^d c/(b^d-a)+f(1)+b^d c/(a-b^d)=f(1)$。归纳步骤：假设对于 k 为真，其中 $n=b^k$。那么对于 $n=b^{k+1}$，$f(n)=af(n/b)+cn^d=a\{[b^d c/(b^d-a)](n/b)^d+[f(1)+b^d c/(a-b^d)]\cdot(n/b)^{\log_b a})\}+cn^d=b^d c/(b^d-a)n^d a/b^d+[f(1)+cb^d c/(a-b^d)]n^{\log_b a}+cn^d=n^d[ac/(b^d-a)+c(b^d-a)/(b^d-a)]+[f(1)+b^d c/(a-b^d c)]n^{\log_b a}=[(b^d c)/(b^d-a)]n^d+[f(1)+b^d c/(a-b^d)]n^{\log_b a}$。

33. 如果 $a>b^d$ 那么 $\log_b a>d$，所以第二项为主，给出 $O(n^{\log_b a})$。

35. $O(n^{\log_4 5})$

37. $O(n^3)$

6.4 节

1. $f(x)=2(x^6-1)/(x-1)$

3. a) $f(x)=2x(1-x^6)/(1-x)$ b) $x^3/(1-x)$ c) $x/(1-x^3)$ d) $2/(1-2x)$ e) $(1+x)^7$ f) $2/(1+x)$ g) $(1/(1-x))-x^2$ h) $x^3/(1-x)^2$

5. a) $5/(1-x)$ b) $1/(1-3x)$ c) $2x^3/(1-x)$ d) $(3-x)/(1-x)^2$ e) $(1+x)^8$ f) $1/(1-x)^5$

7. a) $a_0=-64$，$a_1=144$，$a_2=-108$，$a_3=27$，对于所有的 $n\geqslant 4$ 有 $a_n=0$
   b) 非零系数只有 $a_0=1$，$a_3=3$，$a_6=3$，$a_9=1$
   c) $a_n=5^n$
   d) 对于 $n\geqslant 3$ 有 $a_n=(-3)^{n-3}$，$a_0=a_1=a_2=0$
   e) $a_0=8$，$a_1=3$，$a_2=2$，对于比 2 大的 n，n 为奇数有 $a_n=0$，n 为偶数有 $a_n=1$
   f) 如果 n 是 4 的正整数倍，$a_n=1$；如果 $n<4$ 有 $a_n=-1$；否则 $a_n=0$
   g) 对于 $n\geqslant 2$ 有 $a_n=n-1$，$a_0=a_1=0$
   h) $a_n=2^{n+1}/n!$

9. a) 6 b) 3 c) 9 d) 0 e) 5

11. a) 1024 b) 11 c) 66 d) 292 864 e) 20 412

13. 10

15. 50

17. 20

19. $f(x)=1/((1-x)(1-x^2)(1-x^5)(1-x^{10}))$

21. 15

23. a) $x^4(1+x+x^2+x^3)^2/(1-x)$ b) 6

25. a) 在 $1/((1-x^3)(1-x^4)(1-x^{20}))$ 的幂级数展开式中 x^r 的系数
    b) $1/(1-x^3-x^4-x^{20})$ c) 7 d) 3224

27. a) 3 b) 29 c) 29 d) 242

29. a) 10 b) 49 c) 2 d) 4

31. a) $G(x)-a_0-a_1x-a_2x^2$ b) $G(x^2)$ c) $x^4G(x)$ d) $G(2x)$ e) $\int_0^x G(t)\,dt$ f) $G(x)/(1-x)$

33. $a_k=2\cdot 3^k-1$

35. $a_k=18\cdot 3^k-12\cdot 2^k$

37. $a_k=k^2+8k+20+(6k-18)2^k$

39. 设 $G(x)=\sum_{k=0}^{\infty}f_kx^k$。在对和式的序标移位并将序列求和以后，易得 $G(x)-xG(x)-x^2G(x)=f_0+(f_1-f_0)x+\sum_{k=2}^{\infty}(f_k-f_{k-1}-f_{k-2})x^k=0+x+\sum_{k=2}^{\infty}0x^k$，由此得 $G(x)-xG(x)-x^2G(x)=x$。求解 $G(x)$ 得 $G(x)=x/(1-x-x^2)$。由部分分式的方法证明 $x/(1-x-x^2)=(1/\sqrt{5})[1/(1-\alpha x)-1/(1-\beta x)]$，其中 $\alpha=(1+\sqrt{5})/2$ 且 $\beta=(1-\sqrt{5})/2$。由事实 $1/(1-\alpha x)=\sum_{k=0}^{\infty}\alpha^kx^k$，得 $G(x)=(1/\sqrt{5})\cdot\sum_{k=0}^{\infty}(\alpha^k-\beta^k)x^k$，因此 $f_k=(1/\sqrt{5})(\alpha^k-\beta^k)$。

41. a)设 $G(x)=\sum_{n=0}^{\infty}C_nx^n$ 是 $\{C_n\}$ 的生成函数。那么 $G(x)^2=\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}C_kC_{n-k}\right)x^n=\sum_{n=1}^{\infty}\left(\sum_{k=1}^{n=0}C_kC_{n-1-k}\right)x^{n-1}=\sum_{n=1}^{\infty}C_nx^{n-1}$，因此 $xG(x)^2=\sum_{n=1}^{\infty}C_nx^n$，这就推出 $xG(x)^2-G(x)+1=0$。应用二次方程求根公式得 $G(x)=\dfrac{1\pm\sqrt{1-4x}}{2x}$，在这个公式中我们选择减号，因为选择加号会导致除以零。

b)由练习 40 知 $(1-4x)^{-1/2}=\sum_{n=0}^{\infty}\binom{2n}{n}x^n$，逐项积分（根据微积分的定理是允许的）得 $\int_0^x(1-4t)^{-1/2}dt=\sum_{n=0}^{\infty}\frac{1}{n+1}\binom{2n}{n}x^{n+1}=x\sum_{n=0}^{\infty}\frac{1}{n+1}\binom{2n}{n}x^n$，因为 $\int_0^x(1-4t)^{-1/2}dt=\dfrac{1-\sqrt{1-4x}}{2}=xG(x)$，由系数相等得 $C_n=\dfrac{1}{n+1}\dbinom{2n}{n}$。

43. 对等式 $(1+x)^{m+n}=(1+x)^m(1+x)^n$ 应用二项式定理，证明 $\sum_{r=0}^{m+n}C(m+n,r)x^r=\sum_{r=0}^{m}C(m,r)x^r\cdot\sum_{r=0}C(n,r)x^r=\sum_{r=0}^{m+n}\left(\sum_{k=0}^{r}C(m,r-k)C(n,k)\right)x^r$。比较系数就得到所需要的恒等式。

45. a) $2e^x$ b) e^{-x} c) e^{3x}

d) xe^x+e^x e) $(e^x-1)/x$

47. a) $a_n=(-1)^n$ b) $a_n=3\cdot 2^n$ c) $a_n=3^n-3\cdot 2^n$

d) $a_n=(-2)^n$，$n\geqslant 2$，$a_1=-3$，$a_0=2$

e) $a_n=(-2)^n+n!$

f) $a_n=(-3)^n+n!\cdot 2^n$，$n\geqslant 2$，$a_0=1$，$a_1=-2$

g)若 n 是奇数，$a_n=0$；若 n 是偶数，$a_n=n!/(n/2)!$

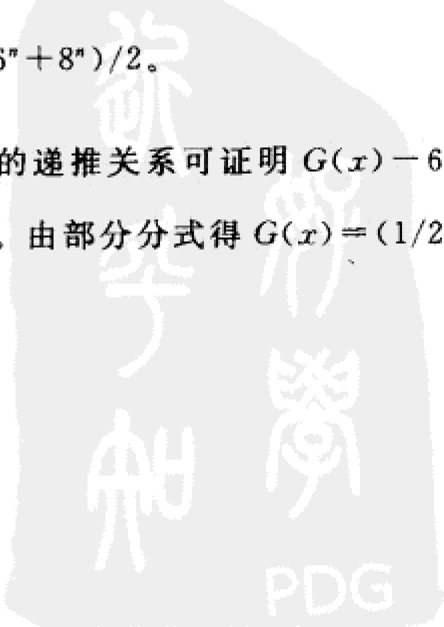
49. a) $a_n=6a_{n-1}+8^{n-1}$，$n\geqslant 1$，$a_0=1$。

b)相关的线性齐次递推关系的通解是 $a_n^{(h)}=\alpha 6^n$，特解是 $a_n^{(p)}=\frac{1}{2}\cdot 8^n$。因此通解是 $a_n=\alpha 6^n+\frac{1}{2}\cdot 8^n$。使用初始条件得 $\alpha=\frac{1}{2}$，因此 $a_n=(6^n+8^n)/2$。

c)设 $G(x)=\sum_{k=0}^{\infty}a_kx^k$。使用关于 $\{a_k\}$ 的递推关系可证明 $G(x)-6xG(x)=(1-7x)/(1-8x)$，因此 $G(x)=(1-7x)/((1-6x)(1-8x))$。由部分分式得 $G(x)=(1/2)/(1-6x)+(1/2)/(1-8x)$。利用表 6-1 得 $a_n=(6^n+8^n)/2$。

51. $\dfrac{1}{1-x}\cdot\dfrac{1}{1-x^2}\cdot\dfrac{1}{1-x^3}\cdots$

53. $(1+x)(1+x)^2(1+x)^3\cdots$



55. 因为$(1+x)(1+x^2)(1+x^3)\cdots=\frac{1-x^2}{1-x}\cdot\frac{1-x^4}{1-x^2}\cdot\frac{1-x^6}{1-x^3}\cdots=\frac{1}{1-x}\cdot\frac{1}{1-x^3}\cdot\frac{1}{1-x^5}\cdots$，练习 52 和 53 的生成函数相等。

57. a) $G_X(1)=\sum_{k=0}^{\infty}p(X=k)\cdot 1^k=\sum_{k=0}^{\infty}P(X=k)=1$

b) $G_X'(1)=\frac{d}{dx}\sum_{k=0}^{\infty}p(X=k)\cdot x^k\,|_{x=1}=\sum_{k=0}^{\infty}p(X=k)\cdot k\cdot x^{k-1}\,|_{x=1}$

$=\sum_{k=0}^{\infty}p(X=k)\cdot k=E(X)$

c) $G_X''(1)=\frac{d^2}{dx^2}\sum_{k=0}^{\infty}p(X=k)\cdot x^k\,|_{x=1}=\sum_{k=0}^{\infty}p(X=k)\cdot k(k-1)\cdot x^{k-2}\,|_{x=1}$

$=\sum_{k=0}^{\infty}p(X=k)\cdot(k^2-k)=V(X)+E(X)^2-E(X)$。与(b) 组合就得到所需的结果

59. a) $G(x)p^m/(1-qx)^m$ b) $V(x)=mq/p^2$

6.5 节

1. a)30 b)29 c)24 d)18
3. 1%
5. a)300 b)150 c)175 d)100
7. 492
9. 974
11. 55
13. 248
15. 50 138
17. 234
19. $|A_1\cup A_2\cup A_3\cup A_4\cup A_5|=|A_1|+|A_2|+|A_3|+|A_4|+|A_5|-|A_1\cap A_2|-|A_1\cap A_3|-|A_1\cap A_4|-|A_1\cap A_5|-|A_2\cap A_3|-|A_2\cap A_4|-|A_2\cap A_5|-|A_3\cap A_4|-|A_3\cap A_5|-|A_4\cap A_5|+|A_1\cap A_2\cap A_3|+|A_1\cap A_2\cap A_4|+|A_1\cap A_2\cap A_5|+|A_1\cap A_3\cap A_4|+|A_1\cap A_3\cap A_5|+|A_1\cap A_4\cap A_5|+|A_2\cap A_3\cap A_4|+|A_2\cap A_3\cap A_5|+|A_2\cap A_4\cap A_5|+|A_3\cap A_4\cap A_5|-|A_1\cap A_2\cap A_3\cap A_4|-|A_1\cap A_2\cap A_3\cap A_5|-|A_1\cap A_2\cap A_4\cap A_5|-|A_1\cap A_3\cap A_4\cap A_5|-|A_2\cap A_3\cap A_4\cap A_5|+|A_1\cap A_2\cap A_3\cap A_4\cap A_5|$
21. $|A_1\cup A_2\cup A_3\cup A_4\cup A_5\cup A_6|=|A_1|+|A_2|+|A_3|+|A_4|+|A_5|+|A_6|-|A_1\cap A_2|-|A_1\cap A_3|-|A_1\cap A_4|-|A_1\cap A_5|-|A_1\cap A_6|-|A_2\cap A_3|-|A_2\cap A_4|-|A_2\cap A_5|-|A_2\cap A_6|-|A_3\cap A_4|-|A_3\cap A_5|-|A_3\cap A_6|-|A_4\cap A_5|-|A_4\cap A_6|-|A_5\cap A_6|$
23. $p(E_1\cup E_2\cup E_3)=p(E_1)+p(E_2)+p(E_3)-p(E_1(E_2)-p(E_1\cap E_3)-p(E_2\cap E_3)+p(E_1\cap E_2\cap E_3)$
25. 4972/71，295
27. $p(E_1\cup E_2\cup E_3\cup E_4\cup E_5)=p(E_1)+p(E_2)+p(E_3)+p(E_4)+p(E_5)-p(E_1\cap E_2)-p(E_1\cap E_3)-p(E_1\cap E_4)-p(E_1\cap E_5)-p(E_2\cap E_3)-p(E_2\cap E_4)-p(E_2\cap E_5)-p(E_3\cap E_4)-p(E_3\cap E_5)-p(E_4\cap E_5)+p(E_1\cap E_2\cap E_3)+p(E_1\cap E_2\cap E_4)+p(E_1\cap E_2\cap E_5)+p(E_1\cap E_3\cap E_4)+p(E_1\cap E_3\cap E_5)+p(E_1\cap E_4\cap E_5)+p(E_2\cap E_3\cap E_4)+p(E_2\cap E_3\cap E_5)+p(E_2\cap E_4\cap E_5)+p(E_3\cap E_4\cap E_5)$
29. $p(\bigcup_{i=1}^{n}E_i)=\sum_{1\leqslant i\leqslant n}p(E_i)-\sum_{1\leqslant i<j\leqslant n}p(E_i\cap E_j)+\sum_{1\leqslant i<j<k\leqslant n}p(E_i\cap E_j\cap E_k)-\cdots+(-1)^{n+1}p(\bigcap_{i=1}^{n}E_i)$

6.6 节

1. 75

3. 6

5. 46

7. 9875

9. 540

11. 2100

13. 1854

15. a) $D_{100}/100!$　　b) $100D_{99}/100!$　　c) $C(100, 2)/100!$　　d) 0　　e) $1/100!$

17. 2 170 680

19. 由练习 18 有 $D_n - nD_{n-1} = -(D_{n-1} - (n-1)D_{n-2})$。通过迭代以及 $D_2 = 1$ 和 $D_1 = 0$，得 $D_n - nD_{n-1} = -(D_{n-1} - (n-1)D_{n-2}) = -(-(D_{n-2} - (n-2)D_{n-3})) = D_{n-2} - (n-2)D_{n-3} = \cdots = (-1)^n(D_2 - 2D_1) = (-1)^n$。

21. 当 n 是奇数

23. $\phi(n) = n - \sum_{i=1}^{m} \frac{n}{p_i} + \sum_{1 \leqslant i < j \leqslant m} \frac{n}{p_i p_j} - \cdots \pm \frac{n}{p_1 p_2 \cdots p_m} = n \prod_{i=1}^{m} (a - \frac{1}{p_i})$

25. 4

27. 从 m 元素集合到 n 元素集合存在 n^m 个函数，从 m 元素集合到 n 元素集合有 $C(n, 1)(n-1)^m$ 个函数恰好缺少 1 个元素，$C(n, 2)(n-2)^m$ 个函数恰好缺少 2 个元素，继续下去，有 $C(n, n-1)1^m$ 个函数恰好缺少 $n-1$ 个元素。因此由容斥原理存在 $n^m - C(n, 1)(n-1)^m + C(n, 2)(n-2)^m - \cdots + (-1)^{n-1} C(n, n-1)1^m$ 个映上的函数。

补充练习

1. a) $A_n = 4A_{n-1}$　　b) $A_1 = 40$　　c) $A_n = 10 \cdot 4^n$

3. a) $M_n = M_{n-1} + 160\,000$　　b) $M_1 = 186\,000$　　c) $M_n = 160\,000n + 26\,000$
   d) $T_n = T_{n-1} + 160\,000n + 26\,000$　　e) $T_n = 80\,000n^2 + 106\,000n$

5. a) $a_n = a_{n-2} + a_{n-3}$　　b) $a_1 = 0,\ a_2 = 1,\ a_3 = 1$　　c) $a_{12} = 12$

7. a) 2　　b) 5　　c) 8　　d) 16

9. $a_n = 2^n$

11. $a_n = 2 + 4n/3 + n^2/2 + n^3/6$

13. $a_n = a_{n-2} + a_{n-3}$

15. $O(n^4)$

17. $O(n)$

19. a) $18n + 18$　　b) 18　　c) 0

21. $\triangle(a_n b_n) = a_{n+1}b_{n+1} - a_n b_n = a_{n+1}(b_{n+1} - b_n) + b_n(a_{n+1} - a_n) = a_{n+1}\triangle b_n + b_n \triangle a_n$

23. a) 设 $G(x) = \sum_{n=0}^{\infty} a_n x^n$，那么 $G'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} = \sum_{n=0}^{\infty} (n+1)a_{n+1}x^n$，因此，$G'(x) - G(x) = \sum_{n=0}^{\infty} ((n+1)a_{n+1} - a_n)x^n = \sum_{n=0} x^n/n! = e^x$，这正是题目要求的。易见 $G(0) = a_0 = 1$。

   b) 我们有 $(e^{-x}G(x))' = e^{-x}G'(x) - e^{-x}G(x) = e^{-x}(G'(x) - G(x)) = e^{-x}e^x = 1$。因此，$e^{-x}G(x) = x + c$，其中 c 是常数。从而，$G(x) = xe^x + ce^x$。由于 $G(0) = 1$，所以 $c = 1$。

   c) 我们有 $G(x) = \sum_{n=0}^{\infty} (x^{n+1}/n!) + \sum_{n=0}^{\infty} (x^n/n!) = \sum_{n=1}^{\infty} (x^n/(n-1)!) + \sum_{n=0}^{\infty} (x^n/n!)$ 由此，对于所有的 $n \geqslant 1$，$a_n = 1/(n-1)! + 1/n!$，且 $a_0 = 1$。

25. 7

27. 110

29.0

31.a)19　b)65　c)122　d)167　e)168

33.$D_{n-1}/(n-1)!$

35.11/32

第7章

7.1节

1. a)$\{(0, 0), (1, 1), (2, 2), (3, 3)\}$

   b)$\{(1, 3), (2, 2), (3, 1), (4, 0)\}$

   c)$\{(1, 0), (2, 0), (2, 1), (3, 0), (3, 1), (3, 2), (4, 0), (4, 1), (4, 2), (4, 3)\}$

   d)$\{(1, 0), (1, 1), (1, 2), (1, 3), (2, 0), (2, 2), (3, 0), (3, 3), (4, 0)\}$(假设0不除以0)

   e)$\{(0, 1), (1, 0), (1, 1), (1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2), (4, 1), (4, 3)\}$

   f)$\{(1, 2), (2, 1), (2, 2)\}$

3. a)传递的　b)自反的，对称的，传递的　c)对称的　d)反对称的

   e)自反的，对称的，反对称的，传递的　f)没有这些性质

5. a)自反的，传递的　b)对称的　c)对称的　d)对称的

7. a)对称的　b)对称的，传递的　c)对称的

   d)自反的，对称的，传递的　e)自反的，传递的　f)自反的，对称的，传递的

   g)反对称的　h)反对称的，传递的

9.c)，d)，f)

11. a)不是反自反的　b)不是反自反的　c)不是反自反的　d)不是反自反的

13. 是，例如在$\{1, 2\}$上的关系$\{(1, 1)\}$

15.$(a, b)\in R$当且仅当a比b高

17.(a)

19. 没有

21. $\forall a\forall b((a, b)\in R\rightarrow(b, a)\notin R)$

23. 2^{mn}

25. a)$\{(a, b)\mid b$整除$a\}$　b)$\{(a, b)\mid a$不整除$b\}$

27. f^{-1}的图

29. a)$\{(a, b)\mid$要求a读书b或者a已经读过书$b\}$

    b)$\{(a, b)\mid$要求a读书b并且a已经读过书$b\}$

    c)$\{(a, b)\mid$或者要求a读书b但是a还没有读过书b，或者不要求a读书b但是a已经读过书$b\}$

    d)$\{(a, b)\mid$要求a读书b但是a还没有读过书$b\}$

    e)$\{(a, b)\mid$不要求a读书b但是a已经读过书$b\}$

31. $S\circ R=\{(a, b)\mid a$是b的父母且b有一个兄弟姐妹$\}$

    $R\circ S=\{(a, b)\mid a$是b的叔伯或阿姨$\}$

33.a)R^2　b)R_6　c)R_3　d)R_3　e)$\varnothing$　f)R_1　g)R_4　h)R_4

35.a)R_1　b)R_2　c)R_3　d)R^2　e)R_3　f)R^2　g)R^2　h)R^2

37. b在某个人指导下得到他或她的博士学位，而这个人在a指导下得到他或她的博士学位；存在一个$n+1$个人的序列，从a开始到b结束，使得序列中每个人都是下一个人的导师。

39. a)$\{(a, b)\mid a-b\equiv 0, 3, 4, 6, 8$或$9 \pmod{12}\}$

    b)$\{(a, b)\mid a\equiv b \pmod{12}\}$

    c)$\{(a, b)\mid a-b\equiv 3, 6$或$9 \pmod{12}\}$

d) $\{(a, b) \mid a-b\equiv 4$ 或 $8(\text{mod } 12)\}$

e) $\{(a, b) \mid a-b\equiv 3, 4, 6, 8$ 或 $9(\text{mod } 12)\}$

41. 8

43. a) 65 536　　b) 32 768

45. a) $2^{n(n+1)/2}$　　b) $2^n3^{n(n-1)/2}$　　c) $3^{n(n-1)/2}$

d) $2^{n(n-1)}$　　e) $2^{n(n-1)/2}$　　f) $2^{n^2}-2\cdot 2^{n(n-1)}$

47. 可能没有这样的 b。

49. 若 R 是对称的且 $(a, b)\in R$，那么 $(b, a)\in R$，因此 $(a, b)\in R^{-1}$，从而有 $R\subseteq R^{-1}$。类似地有 $R^{-1}\subseteq R$，于是 $R=R^{-1}$。相反，若 $R=R^{-1}$ 且 $(a, b)\in R$，那么 $(a, b)\in R^{-1}$，从而 $(b, a)\in R$，于是 R 是对称的。

51. R 是自反的，当且仅当对所有的 $a\in A$ 有 $(a, a)\in R$，当且仅当对所有的 $a\in A$ 有 $(a, a)\in R^{-1}$ [因为 $(a, a)\in R$ 当且仅当 $(a, a)\in R^{-1}$]，从而当且仅当 R^{-1} 是自反的。

53. 使用数学归纳法。对于 $n=1$ 结果是平凡的。假设 R 是自反和传递的。由定理 1，$R^{n+1}\subseteq R$。为得到 $R\subseteq R^{n+1}=R^n\circ R$，令 $(a, b)\in R$。由归纳假设，$R^n=R$ 并且因此是自反的。于是 $(b, b)\in R^n$，从而 $(a, b)\in R^{n+1}$。

55. 使用数学归纳法。$n=1$ 时结果是平凡的。假设 R^n 是自反的，那么对所有的 $a\in A$ 有 $(a, a)\in R^n$ 且 $(a, a)\in R$。于是对所有的 $a\in A$ 有 $(a, a)\in R^n\circ R=R^{n+1}$。

57. 不是，例如 $R=\{(1, 2), (2, 1)\}$。

7.2 节

1. $\{(1, 2, 3), (1, 2, 4), (1, 3, 4), (2, 3, 4)\}$

3. (Nadir，122，34，底特律，08:10)，(Acme，221，22，丹佛，08:17)，(Acme，122，33，安克雷奇，08:22)，(Acme，323，34，檀香山，08:30)，(Nadir，199，13，底特律，08:47)，(Acme，222，22，丹佛，09:10)，(Nadir，322，34，底特律，09:44)

5. 航空公司与航班号，航空公司与起飞时间

7. a)是　　b)不是　　c)不是

9. a)社会保险号

b)没有两个人恰好有着同一街道住址且有相同的名字

c)没有两个有相同的名字的人住在一起

11. (Nadir，122，34，底特律，08:10)，(Nadir，199，13，底特律，08:47)，(Nadir，322，34，底特律，09:44)

13. (Nadir，122，34，底特律，08:10)，(Nadir，199，13，底特律，08:47)，(Nadir，322，34，底特律，09:44)，(Acme，221，22，丹佛，08:17)，(Acme，222，22，丹佛，09:10)

15. $P_{3,5,6}$

17.

航空公司	目的地	航空公司	目的地
Nadir	底特律	Acme	安克雷奇
Acme	丹佛	Acme	檀香山

19.

供货商	零件号	项目	数量	颜色代码
23	1092	1	2	2
23	1101	3	1	1
23	9048	4	12	2

(续)

供货商	零件号	项目	数量	颜色代码
31	4975	3	6	2
31	3477	2	25	2
32	6984	4	10	1
32	9191	2	80	4
33	1001	1	14	8

21. 这个等式的两边选择了由满足 C_1 和 C_2 两个条件的那些 n 元组构成的子集。

23. 这个等式的两边选择了在 R 中，在 S 中，并且满足条件 C 的那些 n 元组构成的子集。

25. 这个等式的两边选择了在 R 或 S 中，由 n 元组的第 i_1，第 i_2，…，第 i_m 分量构成的 m 元组。

27. 设 $R=\{(a, b)\}$ 且 $S=\{(a, c)\}$，$n=2$，$m=1$ 并且 $i_1=1$；$P_1(R-S)=\{(a)\}$，但是 $P_1(R)-P_1(S)=\varnothing$。

29. a) J_2 被 $P_{1,3}$ 跟随着

b) (23, 1), (23, 3), (31, 3) (32, 4)

7.3 节

1. a) $\begin{bmatrix}1&1&1\\0&0&0\\0&0&0\end{bmatrix}$ b) $\begin{bmatrix}0&1&0\\1&1&0\\0&0&1\end{bmatrix}$ c) $\begin{bmatrix}1&1&1\\0&1&1\\0&0&1\end{bmatrix}$ d) $\begin{bmatrix}0&0&1\\0&0&0\\1&0&0\end{bmatrix}$

3. a) (1, 1), (1, 3), (2, 2), (3, 1), (3, 3)　b) (1, 2), (2, 2), (3, 2)

c) (1, 2), (1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2), (3, 3)

5. R 不是反自反的当且仅当在有向图中至少存在一个环。

7. a) 自反的、对称的、传递的

b) 反对称的、传递的

c) 对称的

9. a) 4950　b) 9900　c) 99　d) 100　e) 1

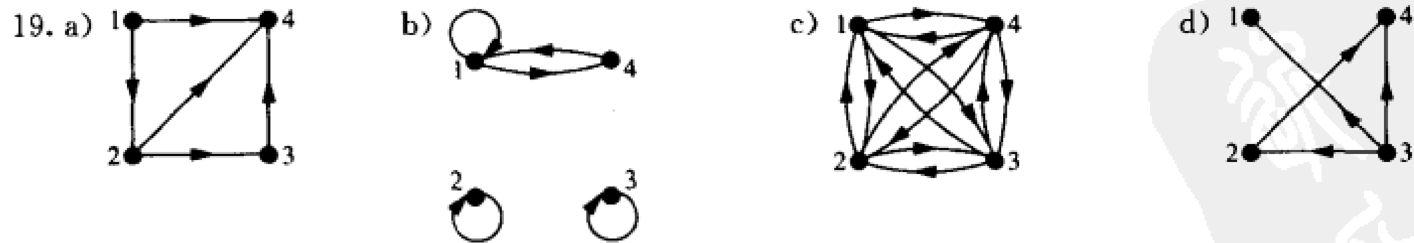
11. 把每个 0 变成 1 并且把每个 1 变成 0。

13. a) $\begin{bmatrix}0&1&1\\1&1&0\\1&0&1\end{bmatrix}$ b) $\begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}$ c) $\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$

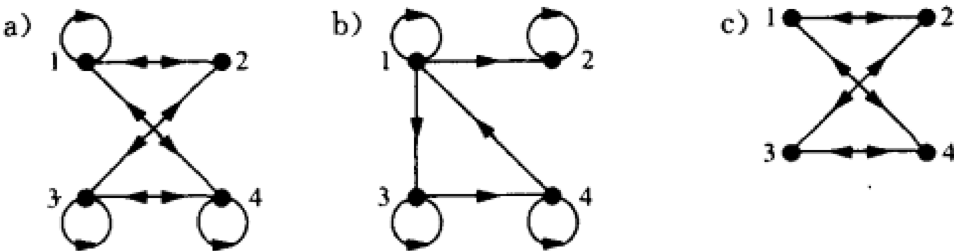
15. a) $\begin{bmatrix}0&0&1\\1&1&0\\0&1&1\end{bmatrix}$ b) $\begin{bmatrix}1&1&0\\0&1&1\\1&1&1\end{bmatrix}$ c) $\begin{bmatrix}0&1&1\\1&1&1\\1&1&1\end{bmatrix}$

17. n^2-k

19.



21. 为了简单起见，我们用双箭头而不是画两条分离的线来表示在同样两个结点之间的方向相反的一对边。



23. {(*a*, *b*), (*a*, *c*), (*b*, *c*), (*c*, *b*)}

25. (*a*, *c*), (*b*, *a*), (*c*, *d*), (*d*, *b*)

27. {(*a*, *a*), (*a*, *b*), (*a*, *c*), (*b*, *a*), (*b*, *b*), (*b*, *c*), (*c*, *a*), (*c*, *b*), (*d*, *d*)}

29. 关系是非对称的，当且仅当有向图没有环，也没有长度 2 的封闭路径。

31. 练习 23：反自反的。练习 24：自反的、反对称的、传递的。练习 25：反自反的、反对称的。

33. 把 R 的有向图的每条边的方向倒转。

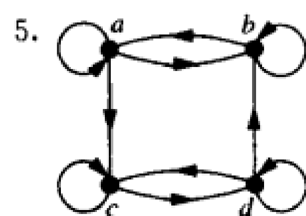
35. 用数学归纳法证明。基础步骤：$n=1$ 是平凡的。归纳步骤：假设对 k 为真。因为 $R^{k+1}=R^k \circ R$，它的矩阵是 $\mathbf{M}_R \odot \mathbf{M}_{R^k}$，由归纳假设就是 $\mathbf{M}_R \odot \mathbf{M}_R^{[k]} = \mathbf{M}_R^{[k+1]}$。

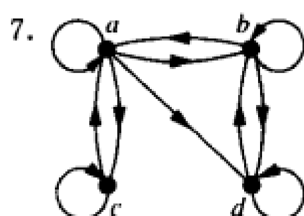
7.4 节

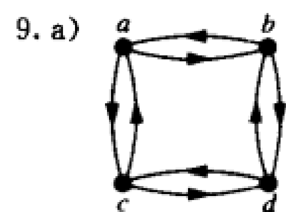
1. a) {(0, 0), (0, 1), (1, 1), (1, 2), (2, 0), (2, 2), (3, 0), (3, 3)}

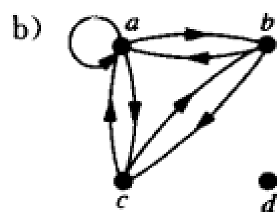
   b) {(0, 1), (0, 2), (0, 3), (1, 0), (1, 1), (1, 2), (2, 0), (2, 1), (2, 2), (3, 0)}

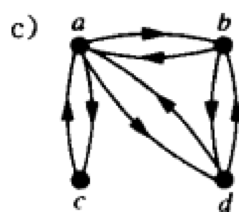
3. {(*a*, *b*) | *a* 整除 *b* 或 *b* 整除 *a*}

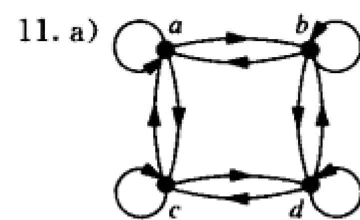
5. 

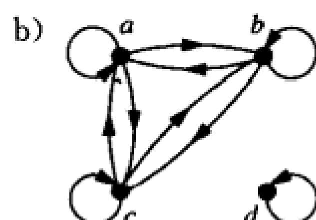
7. 

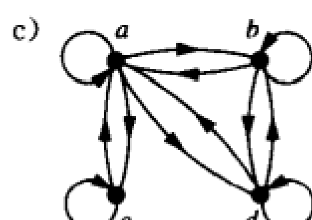
9. a) 

b) 

c) 

11. a) 

b) 

c) 

13. R 的对称闭包是 $R\cup R^{-1}$。$\mathbf{M}_{R\cup R^{-1}}=\mathbf{M}_R \vee \mathbf{M}_{R-1}=\mathbf{M}_R \vee \mathbf{M}_R^t$。

15. 仅当 R 是反自反关系时，它就是它自己的闭包。

17. *a*, *a*, *a*, *a*; *b*, *c*, *c*, *b*; *c*, *b*, *c*, *c*; *c*, *c*, *b*, *c*; *c*, *c*, *c*, *c*; *d*, *e*, *e*, *d*; *e*, *d*, *e*, *e*; *e*, *e*, *d*, *e*; *e*, *e*, *e*, *e*

19. a) {(1, 1), (1, 5), (2, 3), (3, 1), (3, 2), (3, 3), (3, 4), (4, 1), (4, 5), (5, 3), (5, 4)}

    b) {(1, 1), (1, 2), (1, 3), (1, 4), (2, 1), (2, 5), (3, 1), (3, 3), (3, 4), (3, 5), (4, 1), (4, 2), (4, 3), (4, 4), (5, 1), (5, 3), (5, 5)}

    c) {(1, 1)(1, 3), (1, 4), (1, 5), (2, 1), (2, 2), (2, 3), (2, 4), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (4, 1), (4, 3), (4, 4), (4, 5), (5, 1), (5, 2), (5, 3), (5, 4), (5, 5)}

    d) {(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (2, 1), (2, 3), (2, 4), (2, 5), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (5, 1), (5, 2), (5, 3), (5, 4), (5, 5)}

    e) {(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (5, 1), (5, 2), (5, 3), (5, 4), (5, 5)}

    f) {(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (4, 1), (4, 2), (4, 3), (4, 4), (4, 5), (5, 1), (5, 2), (5, 3), (5, 4), (5, 5)}

21. a)如果有学生 c，使得 a 和 c 在同一个班，c 和 b 也在同一个班。
    b)如果有两个学生 c 和 d，使得 a 和 c 在同一个班，且 c 和 d 在同一个班，d 和 b 也在同一个班。
    c)如果存在学生序列 s_0，s_1，…，s_n，$n\geqslant 1$，使得 $s_0=a$，$s_n=b$，且对于每个 $i=1, 2, \cdots, n$，s_i 与 s_{i-1} 在同一个班。

23. 从 $(R^*)^{-1}=\left(\bigcup_{n=1}^{\infty}R^n\right)^{-1}=\bigcup_{n=1}^{\infty}(R^n)^{-1}=\bigcup_{n=1}^{\infty}R^n=R^*$，命题得证。

25. a) $\begin{bmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{bmatrix}$ b) $\begin{bmatrix}0&0&0&0\\1&0&1&1\\1&0&1&1\\1&0&1&1\end{bmatrix}$ c) $\begin{bmatrix}0&1&1&1\\0&0&1&1\\0&0&0&1\\0&0&0&0\end{bmatrix}$ d) $\begin{bmatrix}1&1&1&1\\1&1&1&1\\1&1&1&1\\1&1&1&1\end{bmatrix}$

27. 答案与练习 25 一样。

29. a) $\{(1, 1), (1, 2), (1, 4), (2, 2), (3, 3), (4, 1), (4, 2), (4, 4)\}$
    b) $\{(1, 1), (1, 2), (1, 4), (2, 1), (2, 2), (2, 4), (3, 3), (4, 1), (4, 2), (4, 4)\}$
    c) $\{(1, 1), (1, 2), (1, 4), (2, 1), (2, 2), (2, 4), (3, 3), (4, 1), (4, 2), (4, 4)\}$

31. 算法 1：$O(n^{3.8})$；算法 2：$O(n^3)$

33. 初始令 $\mathbf{A}:=\mathbf{M}_R\vee\mathbf{I}_n$ 且只对 $i:=2$ 到 $n-1$ 循环。

35. a)因为 R 是自反的，每个包含它的关系也一定是自反的。
    b) $\{(0, 0), (0, 1), (0, 2), (1, 1), (2, 2)\}$ 和 $\{(0, 0), (0, 1), (1, 0), (1, 1), (2, 2)\}$ 都包含 R 且都有奇数个元素，但是其中的每一个都不是另一个的子集。

7.5 节

1. a)等价关系 b)不是自反的，也不是传递的 c)等价关系
   d)不是传递的 e)不是对称的，也不是传递的

3. a)等价关系 b)不是传递的 c)不是自反的，不是对称的，也不是传递的
   d)等价关系 e)不是自反的，也不是传递的

5. a)因为 $f(x)=f(x)$，$(x, x)\in R$。所以 R 是自反的。$(x, y)\in R$ 当且仅当 $f(x)=f(y)$，当且仅当 $f(y)=f(x)$，当且仅当 $(y, x)\in R$。因此 R 是对称的。如果 $(x, y)\in R$ 且 $(y, z)\in R$，那么 $f(x)=f(y)$ 且 $f(y)=f(z)$，因此 $f(x)=f(z)$。于是 $(x, z)\in R$，这就证明 R 是传递的。
   b)对于 f 的值域中的 b 得到的集合 $f^{-1}(b)$。

7. 设 x 是长度至少为 3 的串，由于 x 与自己在前 3 位相同，$(x, x)\in R$，因此 R 是自反的。假设 $(x, y)\in R$，那么 x 与 y 在前 3 位相同，因此 y 与 x 也在前 3 位相同，于是 $(y, x)\in R$。如果 (x, y) 和 (y, z) 在 R 中，那么 x 和 y 在前 3 位相同，y 和 z 也在前 3 位相同。因此 x 和 z 在前 3 位相同。从而 $(x, z)\in R$。于是 R 是传递的。

9. 命题 p 等价于 q 意味着 p 和 q 在它们的真值表中有相同的项。R 是自反的，因为 p 与 p 的真值表相同。R 是对称的，因为 p 与 q 有相同的真值表，那么 q 与 p 也有相同的真值表。如果 p 与 q 在它们的真值表中有相同的项并且 q 与 r 在它们的真值表中有相同的项，那么 p 与 r 在它们的真值表中也有相同的项，因此 R 是传递的。

11. a)由练习 5 得出。其中从($\mathbf{R}$ 到 $\mathbf{R}$ 的)可微分函数集合到($\mathbf{R}$ 到 $\mathbf{R}$ 的)函数集合的函数 f 是微分运算。
    b)对某个常数 C 形为 $g(x)=x^2+C$ 的所有函数的集合。

13. a)由练习 5 得出。其中 f 是从所有 URL 的集合到所有 Web 页集合的函数，f 对每个 URL 的赋值是关于这个 URL 的 Web 页。

15. 不是

17. 不是

19. R 是自反的，因为二进制串 s 和它自己有相同数量的 1。R 是对称的，因为 s 和 t 有相同数量的 1 推出 t 和 s 也有相同数量的 1。R 是传递的，因为如果 s 与 t 有相同数量的 1，t 与 u 有相同数量的 1，那么 s 与 u 也有相同数量的 1。

21. a)相同年龄的人的集合

    b)具有相同父母的人的集合

23. 所有恰好具有 2 个 1 的二进制串的集合

25. a)$\{s \mid s$ 是 3 位二进制串$\}$

    b)$\{s1 \mid s$ 是 3 位二进制串$\}$

    c)$\{s11 \mid s$ 是 3 位二进制串$\}$

    d)$\{s1010 \mid s$ 是 3 位二进制串$\}$

27. $\{6n+k \mid n\in \mathbf{Z}\}$，$k\in\{0, 1, 2, 3, 4, 5\}$

29. a)不是　b)是　c)是　d)不是

31. (a)，(c)，(e)

33. (b)，(d)，(e)

35. a)$\{(0, 0), (1, 1), (1, 2), (2, 1), (2, 2), (3, 3), (3, 4), (3, 5), (4, 3), (4, 4), (4, 5), (5, 3), (5, 4), (5, 5)\}$

    b)$\{(0, 0), (0, 1), (1, 0), (1, 1), (2, 2), (2, 3), (3, 2), (3, 3), (4, 4), (4, 5), (5, 4), (5, 5)\}$

    c)$\{(0, 0), (0, 1), (0, 2), (1, 0), (1, 1), (1, 2), (2, 0), (2, 1), (2, 2), (3, 3), (3, 4), (3, 5), (4, 3), (4, 4), (4, 5), (5, 3), (5, 4), (5, 5)\}$

    d)$\{(0, 0), (1, 1), (2, 2), (3, 3), (4, 4), (5, 5)\}$

37. $[0]_6\subseteq[0]_3$，$[1]_6\subseteq[1]_3$，$[2]_6\subseteq[2]_3$，$[3]_6\subseteq[0]_3$，$[4]_6\subseteq[1]_3$，$[5]_6\subseteq[2]_3$

39. 设 A 是在第一个划分中的集合。取 A 中一个特定元素 x。所有 16 位长且与 x 在后 8 位相同的二进制串的集合是在第二个划分中的一个集合。显然，A 中的每个串都在这个集合里。

41. $\{(a, a), (a, b), (a, c), (b, a), (b, b), (b, c), (c, a), (c, b), (c, c), (d, d), (d, e), (e, d), (e, e)\}$

43. a)$\mathbf{Z}$　b)$\{n+\frac{1}{2} \mid n\in\mathbf{Z}\}$

45. a)R 是自反的，因为任何涂色方案通过旋转 360 度可以从自身得到。由于每个旋转是两个翻转的合成，相反，两个翻转的合成也是一个旋转。由这个事实可以说明 R 是对称和传递的。(C_1, C_2)属于 R 当且仅当 C_2 可以通过翻转的合成由 C_1 得到。如果(C_1, C_2)属于 R，那么由于翻转合成的逆也是翻转的合成(按照相反的次序)，(C_2, C_1)也属于 R。于是 R 是对称的。假设(C_1, C_2)和(C_2, C_3)都属于 R，取每种情况下翻转的合成得到的仍是翻转的合成，故(C_1, C_3)属于 R，R 是传递的。

    b)我们用长度为 4 的序列表示涂色，r 和 b 分别代表红和蓝。我们按照左上方格、右上方格、左下方格、右下方格的次序列出表示颜色的字母。等价类是：$\{rrrr\}$，$\{bbbb\}$，$\{rrrb, rrbr, rbrr, brrr\}$，$\{bbbr, bbrb, brbb, rbbb\}$，$\{rbbr, brrb\}$，$\{rrbb, brbr, bbrr, rbrb\}$

47. 5

49. 是

51. R

53. 首先构成 R 的自反闭包，然后构成这个自反闭包的对称闭包，最后构成自反闭包的对称闭包的传递闭包。

55. $p(0)=1$，$p(1)=1$，$p(2)=2$，$p(3)=5$，$p(4)=15$，$p(5)=52$，$p(6)=203$，$p(7)=877$，$p(8)=4140$，$p(9)=21\,147$，$p(10)=115\,975$

7.6 节

1. a)是　　b)不是　　c)是　　d)不是

3. 不是

5. 是

7. a)$\{(0, 0), (1, 0), (1, 1), (2, 0), (2, 1), (2, 2)\}$
   b)$(\mathbf{Z}, \leqslant)$　　c)$(P(\mathbf{Z}), \subseteq)$　　d)$(\mathbf{Z}^+$，"是倍数")

9. a)例如$\{0\}$和$\{1\}$　　b)例如 4 和 6

11. a)$(1, 1, 2) < (1, 2, 1)$　　b)$(0, 1, 2, 3) < (0, 1, 3, 2)$
    c)$(0, 1, 1, 1, 0) < (1, 0, 1, 0, 1)$

13. $0 < 0001 < 001 < 01 < 010 < 0101 < 011 < 11$

15.

17. a)

b)

c)

d)

19. (a, b)，(a, c)，(a, d)，(b, c)，(b, d)，(a, a)，(b, b)，(c, c)，(d, d)

21. (a, a)，(a, g)，(a, d)，(a, e)，(a, f)，(b, b)，(b, g)，(b, d)，(b, e)，(b, f)，(c, c)，(c, g)，(c, d)，(c, e)，(c, f)，(g, d)，(g, e)，(g, f)，(g, g)，(d, d)，(e, e)，(f, f)

23. $(\varnothing, \{a\})$，$(\varnothing, \{b\})$，$(\varnothing, \{c\})$，$(\{a\}, \{a, b\})$，$(\{a\}, \{a, c\})$，$(\{b\}, \{a, b\})$，$(\{b\}, \{b, c\})$，$(\{c\}, \{a, c\})$，$(\{c\}, \{b, c\})$，$(\{a, b\}, \{a, b, c\})$，$(\{a, c\}, \{a, b, c\})$，$(\{b, c\}, \{a, b, c\})$

25. 设$(S, \preccurlyeq)$是有穷偏序集。我们将证明这个偏序集是它的覆盖关系的自反传递闭包。假设(a, b)属于覆盖关系的自反传递闭包。那么$a=b$或$a \prec b$，因此$a \preccurlyeq b$；或者存在一个序列$a_1, a_2, \cdots, a_n$使得$a_1 \prec a_2 \prec \cdots \prec a_n \prec b$，在这种情况下由$\preccurlyeq$的传递性也有$a \preccurlyeq b$。反之，假设$a \prec b$，如果$a=b$，那么$(a, b)$属于覆盖关系的自反传递闭包。如果$a \prec b$并且不存在$z$使得$a \prec z \prec b$，那么$(a, b)$属于覆盖关系并且因此属于它的自反传递闭包。否则，设$a_1 \prec a_2 \prec \cdots \prec a_n \prec b$是这种形式的最长的序列(由于偏序集是有穷的所以存在这样的序列)，那么中间不会有其他元素插入，因此每个对(a, a_1)，(a_1, a_2)，$\cdots$，(a_n, b)属于覆盖关系，从而(a, b)也属于它的自反传递闭包。

27. a)24，45　b)3，5　c)不存在　d)不存在
e)15，45　f)15　g)15，5，3　h)15

29. a){1，2}，{1，3，4}，{2，3，4}　b){1}，{2}，{4}
c)不存在　d)不存在　e){2，4}，{2，3，4}　f){2，4}
g){3，4}，{4}　h){3，4}

31. 因为$(a, b) \preccurlyeq (a, b)$，$\preccurlyeq$是自反的。如果$(a_1, a_2) \preccurlyeq (b_1, b_2)$并且$(a_1, a_2) \neq (b_1, b_2)$，那么或者$a_1 \prec a_2$，或者$a_1 = b_1$并且$a_2 \prec b_2$。在两种情况下$(b_1, b_2)$都不小于等于$(a_1, a_2)$。因此$\preccurlyeq$是反对称的。假设$(a_1, a_2) \prec (b_1, b_2) \prec (c_1, c_2)$。那么如果$a_1 \prec b_1$或$b_1 \prec c_1$，我们有$a_1 \prec c_1$，所以$(a_1, a_2) \prec (c_1, c_2)$。但是如果$a_1 = b_1 = c_1$，那么$a_2 \prec b_2 \prec c_2$，这也推出$(a_1, a_2) \prec (c_1, c_2)$，于是$\preccurlyeq$是传递的。

33. 因为$(s, t) \preccurlyeq (s, t)$，$\preccurlyeq$是自反的。如果$(s, t) \preccurlyeq (u, v)$并且$(u, v) \preccurlyeq (s, t)$，那么$s \preccurlyeq u \preccurlyeq s$且$t \preccurlyeq v \preccurlyeq t$；从而$s = u$且$t = v$。因此$\preccurlyeq$是反对称的。假设$(s, t) \preccurlyeq (u, v) \preccurlyeq (w, x)$，那么$s \preccurlyeq u$，$t \preccurlyeq v$，$u \preccurlyeq w$，$v \preccurlyeq x$，从而有$s \preccurlyeq w$和$t \preccurlyeq x$，因此$(s, t) \preccurlyeq (w, x)$。于是$\preccurlyeq$是传递的。

35. a)假设x是极大元素并且y是最大元素，那么$x \preccurlyeq y$。因为x不小于y，从而$x = y$。由练习32a)y是唯一的。因此x也是唯一的。
b)假设x是极小元素并且y是最小元素，那么$x \succcurlyeq y$。由于x不大于y，从而$x = y$。由练习32b)y是唯一的。因此x也是唯一的。

37. a)是　b)不是　c)是

39. 使用数学归纳法。设$P(n)$是“格的每个n元子集有最小上界和最大下界”。
基础步骤：$P(1)$为真，因为$\{x\}$的最小上界和最大下界都是x。
归纳步骤：假设$P(k)$为真。令S是$n+1$元素集合。设$x \in S$并且$S' = S - \{x\}$。因为S'有n个元素，由归纳假设有最小上界y和最大下界a。由于这是一个格，存在元素$z = \mathrm{lub}(x, y)$和$b = \mathrm{glb}(x, a)$。如果我们可以证明z是S的最小上界和b是S的最大下界，那么命题得证。为证明z是S的最小上界，首先注意到如果$w \in S$，那么$w = x$或$w \in S'$。如果$w = x$，因为z是x和y的最小上界，那么有$w \preccurlyeq z$。如果$w \in S'$，由于y是S'的最小上界，故$w \preccurlyeq y$，且由于$z = \mathrm{lub}(x, y)$，有$y \preccurlyeq z$，因而$w \preccurlyeq z$。为证明z是S的最小上界，假设u是S的一个上界，那么元素u一定是x和y的一个上界，但是由于$z = \mathrm{lub}(x, y)$，从而有$z \preccurlyeq u$。类似的论述可以证明b是S的最大下界。

41. a)不允许　b)允许
c)(私有的，{猎豹，美洲狮})，(受限制的，{猎豹，美洲狮})，(注册的，{猎豹，美洲狮})，(私有的，{猎豹，美洲狮，黑斑羚})，(受限制的，{猎豹，美洲狮，黑斑羚})，(注册的，{猎豹，美洲狮，黑斑羚})
d)(非私有的，{黑斑羚，美洲狮})，(私有的，{黑斑羚，美洲狮})，(受限制的，{黑斑羚，美洲狮})，(非私有的，{黑斑羚})，(私有的，{黑斑羚})，(受限制的，{黑斑羚})，(非私有的，{美洲狮})，(私有的，{美洲狮})，(受限制的，{美洲狮})，(非私有的，$\varnothing$)，(私有的，$\varnothing$)，(受限制的，$\varnothing$)

43. 设Π是集合S的所有划分的集合，如果划分P_1是P_2的加细，即如果P_1中的每个集合都是P_2中某个集合的子集，则$P_1 \preccurlyeq P_2$。首先我们证明$(\Pi, \preccurlyeq)$是偏序集。由于对每个划分P有$P \preccurlyeq P$，$\preccurlyeq$是自反的。现在假设$P_1 \preccurlyeq P_2$并且$P_2 \preccurlyeq P_1$，令$T \in P_1$，因为$P_1 \preccurlyeq P_2$，存在集合$T' \in P_2$使得$T \subseteq T'$；又因为$P_2 \preccurlyeq P_1$，存在集合$T'' \in P_1$使得$T' \subseteq T''$，从而$T \subseteq T''$。但是因为P_1是划分，由$T = T''$和$T \subseteq T' \subseteq T''$推出$T = T'$，于是$T \in P_2$。反之，通过交换$P_1$与$P_2$同样得出$P_2$的每个子集也在$P_1$中。因此$P_1 = P_2$并且$\preccurlyeq$是反对称的。下面假设$P_1 \preccurlyeq P_2$并且$P_2 \preccurlyeq P_3$。设$T \in P_1$，那么存在集合$T' \in P_2$使得$T \subseteq T'$。由于$P_2 \preccurlyeq P_3$，存在集合$T'' \in P_3$使得$T' \subseteq T''$，从而有$T \subseteq T''$，因此$P_1 \preccurlyeq P_3$。即$\preccurlyeq$是传递的。划分$P_1$和$P_2$的最大下界是划分$P$，$P$的子集都是形如$T_1 \cap T_2$的非空集合，其中$T_1 \in P_1$且$T_2 \in P_2$。关于这个结论的理由不再赘述。划分$P_1$与$P_2$的最小上界是对应于下述等价关系的划分。$x \in S$关系到$y \in S$，如果对某个非负整数$n$存在序列$x = x_0, x_1, x_2, \cdots, x_n = y$，使得对从1到$n$的每个$i$，$x_{i-1}$和$x_i$在

P_1 或者在 P_2 的同一个元素中。我们不再详细证明这是一个等价关系，也不再证明它就是两个划分的最小上界。

45. 由练习 39，对整个有穷格存在最小上界和最大下界。根据定义，这些元素分别是最大元素和最小元素。

47. $\mathbf{Z}^+ \times \mathbf{Z}^+$ 的子集的最小元素是具有最小的横坐标的对，并且如果存在多个这样的对，它就是这些对中具有最小的纵坐标的对。

49. 如果 x 是这个偏序集元素的一个递减序列中的整数，那么在这个序列中至多 $|x|$ 个元素可能跟随着 x，即绝对值为 $|x|-1$，$|x|-2$，…，1，0 的整数。因此不可能存在无穷递减序列。这不是一个全序集，例如 5 和 -5 是不可比的。

51. 为找出两个有理数中哪一个更大，把它们写成具有相同正分母的分数并且比较它们的分子。为证明这个集合是稠密的，假设 $x<y$ 是两个有理数。那么它们的平均值，即 $(x+y)/2$ 是一个在它们之间的有理数。

53. 设 $\langle S, \preccurlyeq \rangle$ 是偏序集，只需证明 S 的每个非空子集包含一个最小元素，当且仅当在 S 中不存在元素的无限递减序列 a_1，a_2，a_3，…(即对所有的 i，$a_{i+1} \prec a_i$)。一个元素的无限递减序列显然没有最小元素。反之，设 A 是 S 的一个没有最小元素的非空子集。由于 A 非空，选择 $a_1 \in A$。因为 a_1 不是 A 的最小元素，选择 $a_2 \in A$ 且 $a_2 \prec a_1$。因为 a_2 不是 A 的最小元素，选择 $a_3 \in A$ 且 $a_3 \prec a_2$。按这种方法继续做下去，产生一个 S 中的无限递减序列。

55. $a \prec b \prec c \prec d \prec e \prec f \prec g \prec h \prec i \prec j \prec k \prec l \prec m$

57. $C \prec A \prec B \prec D \prec E \prec F \prec G$

59. 确定用户需求 $\prec$ 写出功能需求 $\prec$ 设置测试点 $\prec$ 开发系统需求 $\prec$ 写文档 $\prec$ 开发模块 $A \prec$ 开发模块 $B \prec$ 开发模块 $C \prec$ 模块集成 $\prec \alpha$ 测试 $\prec \beta$ 测试 $\prec$ 完成

补充练习

1. a)反自反的(这里不包含空串)，对称的
   b)反自反的，对称的
   c)反自反的，反对称的，传递的

3. 因为 $a+b=a+b$，$((a, b), (a, b)) \in R$，因此 R 是自反的。如果 $((a, b), (c, d)) \in R$ 那么 $a+d=b+c$，因此 $c+b=d+a$。从而 $((c, d), (a, b)) \in R$，于是 R 是对称的。假设 $((a, b), (c, d))$ 和 $((c, d), (e, f))$ 属于 R，那么 $a+d=b+c$ 且 $c+f=d+e$。把这两个等式相加，然后从两边减去 $c+d$ 得到 $a+f=b+e$。从而 $((a, b), (e, f))$ 属于 R，于是 R 是传递的。

5. 假设 $(a, b) \in R$。因为 $(b, b) \in R$，从而有 $(a, b) \in R^2$。

7. 是，是

9. 是，是

11. 在投影中具有相同关键字的两个记录在原来的关系中有相同的关键字。

13. $(\Delta \cup R)^{-1} = \Delta^{-1} \cup R^{-1} = \Delta \cup R^{-1}$

15. a)$R=\{(a, b), (a, c)\}$。R 的对称闭包的传递闭包是 $\{(a, a), (a, b), (a, c), (b, a), (b, b), (b, c), (c, a), (c, b), (c, c)\}$，它与 R 的传递闭包的对称闭包是不同的，后者是 $\{(a, b), (a, c), (b, a), (c, a)\}$。
    b)假设 (a, b) 属于 R 的传递闭包的对称闭包。必须证明 (a, b) 属于 R 的对称闭包的传递闭包。我们知道 (a, b) 和 (b, a) 中至少有一个属于 R 的传递闭包，因此在 R 中存在一条从 a 到 b 的或从 b 到 a 的路径(或者两条都有)。在前面的情况下，在 R 的对称闭包中存在一条从 a 到 b 的路径。在后面的情况下，我们可以通过把从 b 到 a 路径中的所有的边改变方向向回走，在 R 的对称闭包中构造一条从 a 到 b 的路径。因此 (a, b) 属于 R 的对称闭包的传递闭包。

17. 由于 $R \subseteq S$，S 的关于性质 P 的闭包是包含 R 的具有性质 P 的关系。因此 S 的关于性质 P 的闭包包含

了R关于性质P的闭包。

19. 使用沃舍尔(Warshell)算法的基本思想，不同的只是令$w_{ij}^{[k]}$等于使用下标不超过k的内部顶点从u_i到u_j的最长路径的长度。并且如果没有这种路径令$w_{ij}^{[k]}$等于-1。为了从w_{k-1}的元素找到$w_{ij}^{[k]}$，对于每个(i, j)对确定是否存在不使用序标大于k的顶点从v_i到v_k的路径和从v_k到v_j的路径。如果$w_{ik}^{[k-1]}$或$w_{kj}^{[k-1]}$是-1，那么这对路径不存在，因此置$w_{ij}^{[k]}=w_{ij}^{[k-1]}$。如果这对路径存在，那么存在两种可能。如果$w_{kk}^{[k-1]}>0$，那么存在任意长的从$v_i$到$v_j$的路径，因此置$w_{ij}^{[k]}=\infty$。如果$w_{kk}^{[k-1]}=0$，置$w_{ij}^{-1}=\max(w_{ij}^{[k-1]}, w_{ik}^{[k-1]}+w_{kj}^{[k-1]})$(初始取$\mathbf{W}_0=\mathbf{M}_R$)。

21. 25

23. 因为$A_i\cap B_j$是A_i和B_j的子集，这些子集的集合是每个给定划分的加细。我们必须证明这是一个划分。根据构造，每个这样的集合是非空的。下面说明这些集合的并就是S。假设$s\in S$，因为P_1和P_2是S的划分，存在集合A_i和B_j使得$s\in A_i$且$s\in B_j$，从而$s\in A_i\cap B_j$。因此这些集合的并就是S。为说明它们是两两不相交的，只要注意到除非$i=i'$，$j=j'$，$(A_i\cap B_j)\cap(A_{i'}\cap B_{j'})=(A_i\cap A_{i'})\cap(A_j\cap A_{j'})=\varnothing$。

25. 子集关系是任何集合族上的偏序，因为它是自反的，反对称的，传递的，这里的集合族是$R(S)$。

27. 找菜谱<买海鲜<买杂货<洗海贝<切姜蒜<洗鱼<蒸饭<切鱼<洗蔬菜<切水栗子<配菜<烹调<摆放餐具<上菜

29. a)具有多个元素的唯一的反链是$\{c, d\}$。

    b)具有多个元素的反链只有$\{b, c\}$，$\{c, e\}$和$\{d, e\}$。

    c)具有多个元素的反链只有$\{a, b\}$，$\{a, c\}$，$\{b, c\}$，$\{a, b, c\}$，$\{d, e\}$，$\{d, f\}$，$\{e, f\}$，$\{d, e, f\}$。

31. 设$(S, \preccurlyeq)$是有穷偏序集，且A是一条极大链。因为$(A, \preccurlyeq)$也是一个偏序集，它一定有极小元素m。假设m不是S中的极小元素，那么存在S中的元素a使得$a\prec m$。但是，这就使得集合$A\cup\{a\}$变成比A更大的链，为此我们需要证明a与A中每一个元素都可比。因为m与A中每个元素都可比，并且m是极小元素，当x属于A且$x\neq m$时有$m\prec x$。由于$a\prec m$且$m\prec x$，由传递性对于A中的每个元素x有$a\prec x$。

33. 令aRb表示a是b的后代。由练习32，如果不存在$n+1$个人的集合使得其中每个人都不是其他人的后代(一条反链)，那么$k\leqslant n$。因此这个集合可以被划分成$k\leqslant n$条链。根据鸽巢原理，这些链中至少有一条链包含至少$m+1$个人。

35. 我们通过反证法证明，如果S没有无限递减序列，且$\forall x((\forall y(y\prec x\to P(y)))\to P(x))$，那么$P(x)$对所有的$x\in S$为真。如果$P(x)$为真不是对所有的$x\in S$成立，设$x_1$是一个使得$P(x_1)$不为真的$S$的元素。那么根据已知的提示，一定是$\forall y(y\prec x_1\to P(y))$不为真。这意味着存在某个$x_2(x_2\prec x_1)$使得$P(x_2)$不为真。再次使用这个提示，我们得到$x_3\prec x_2$使得$P(x_3)$不为真，并且可以永远这样做下去。这与偏序集的良基性相矛盾。因此$P(x)$对所有的$x\in S$为真。

37. 假设R是近似序。因为R是自反的，如果$a\in A$，那么有$(a, a)\in R$，从而有$(a, a)\in R^{-1}$。因此$a\in R\cap R^{-1}$。于是$R\cap R^{-1}$是自反的。对任何关系R，$R\cap R^{-1}$是对称的，因为对任何关系R，如果$(a, b)\in R$那么有$(b, a)\in R^{-1}$；反之也对。为证明$R\cap R^{-1}$是传递的，假设$(a, b)\in R\cap R^{-1}$并且$(b, c)\in R\cap R^{-1}$。因为R是传递的，由$(a, b)\in R$和$(b, c)\in R$推出$(a, c)\in R$。类似地，由$(a, b)\in R^{-1}$和$(b, c)\in R^{-1}$，就有$(b, a)\in R$和$(c, b)\in R$，因此$(c, a)\in R$，从而$(a, c)\in R^{-1}$。于是$(a, c)\in R\cap R^{-1}$。从而证明了$R\cap R^{-1}$是等价关系。

39. a)因为$\mathrm{glb}(x, y)=\mathrm{glb}(y, x)$和$\mathrm{lub}(x, y)=\mathrm{lub}(y, x)$，从而有$x\wedge y=y\wedge x$和$x\vee y=y\vee x$。

    b)使用定义，$(x\wedge y)\wedge z$是x、y与z的下界，且大于每一个其他的下界。因为x、y和z是可互换的，$x\wedge(y\wedge z)$是同一元素。类似地，$(x\vee y)\vee z$是x、y与z的上界并且小于每一个其他的上界。因为x、y和z是可互换的，$x\vee(y\vee z)$也是同一元素。

c)为证明 $x\wedge(x\vee y)=x$，只需证明 x 是 x 与 $x\vee y$ 的最大下界。注意到 x 是 x 的下界，根据定义 $x\vee y$ 大于 x，x 也是它的下界。所以 x 是 x 与 $x\vee y$ 的下界。但是 x 的任何下界必须小于或等于 x，因此 x 是最大的下界。第二个公式是第一个公式的对偶式，证明省略。

d)x 是自己与自己的下界，也是自己与自己的上界，并且也是其中的最大下界，最小上界。

41. a)因为 1 是大于或等于 1 的唯一元素，因此它也是 1 的唯一上界，因而也是 x 和 1 最小上界的唯一可能的值。

b)因为 $x\leqslant 1$，x 是 x 和 1 的下界并且没有其他的下界可能大于 x，因此 $x\wedge 1=x$。

c)因为 $0\leqslant x$，x 是 x 和 0 的上界并且没有其他的上界可能小于 x，因此 $x\vee 0=x$。

d)因为 0 是小于或等于 0 的唯一元素，因此它也是 0 的唯一下界，因而也是 x 和 0 最大下界的唯一可能的值。

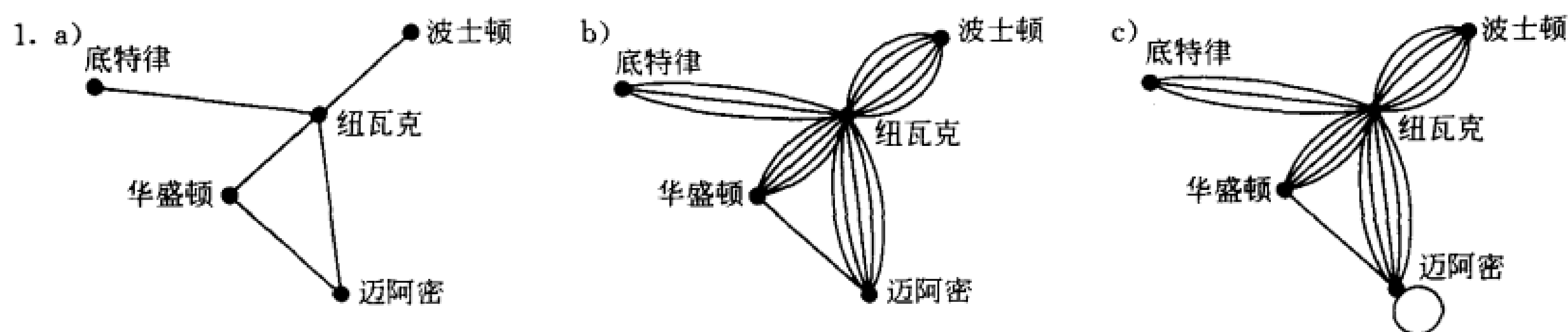
43. $L=(S,\subseteq)$，其中 $S=\{\varnothing,\{1\},\{2\},\{3\},\{1,2\},\{2,3\},\{1,2,3\}\}$

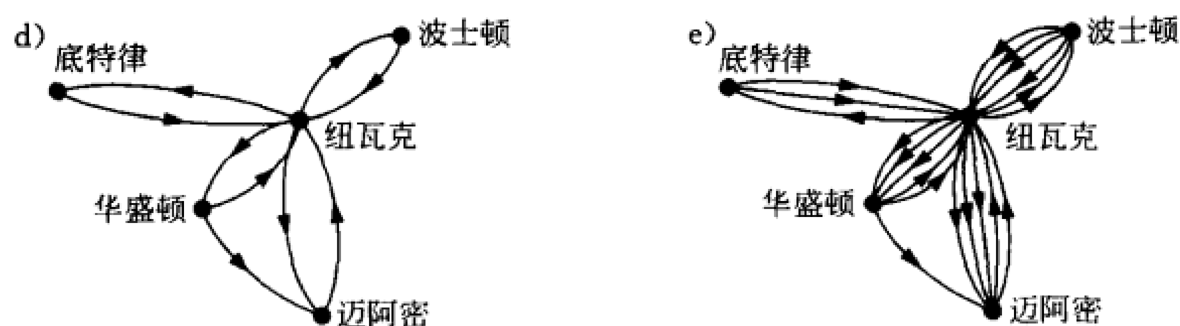
45. 是

47. 子集 $X\subseteq S$ 的补是它的补集 $S-X$。为证明这一点，因为 $X\cup(S-X)=S$ 和 $X\cap(S-X)=\varnothing$，因此 $X\vee(S-X)=1$ 且 $X\wedge(S-X)=0$。

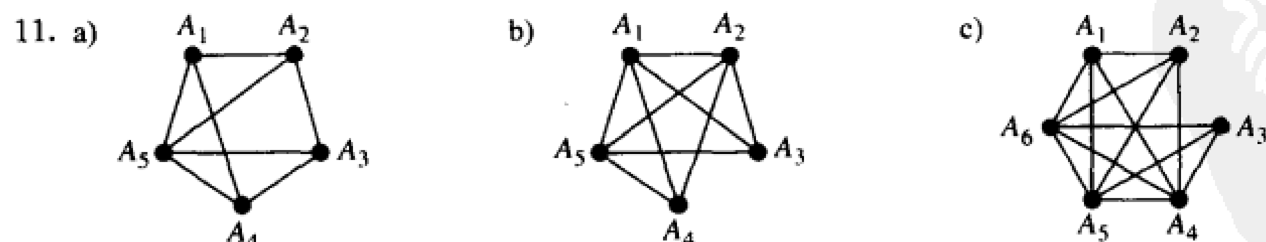
第 8 章

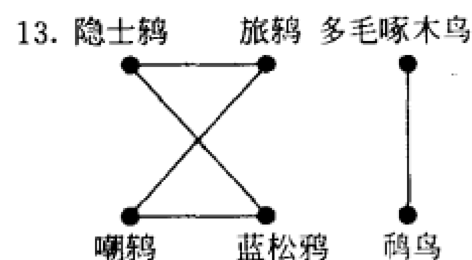
8.1 节



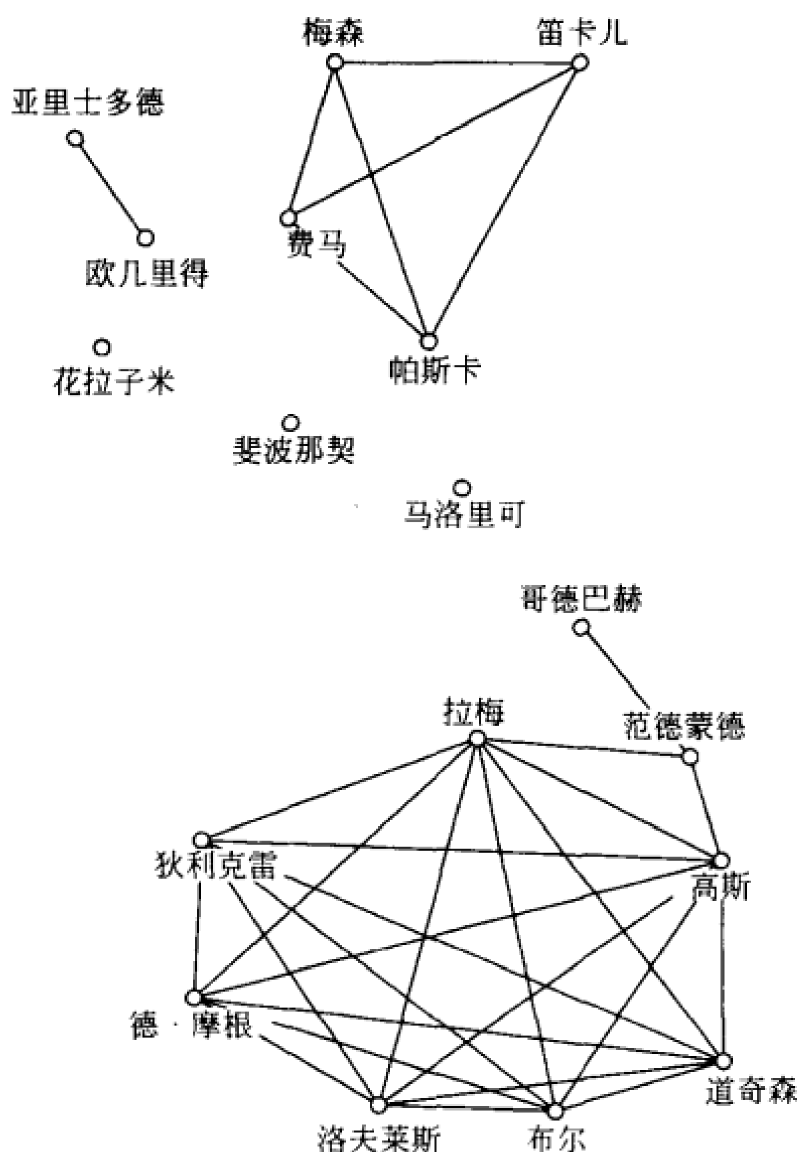


3. 简单图　5. 伪图　7. 有向图　9. 有向多重图

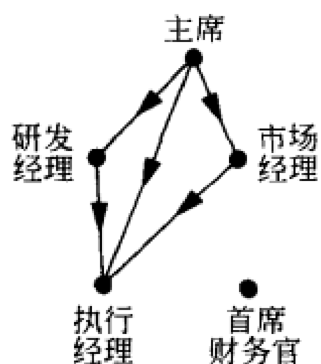




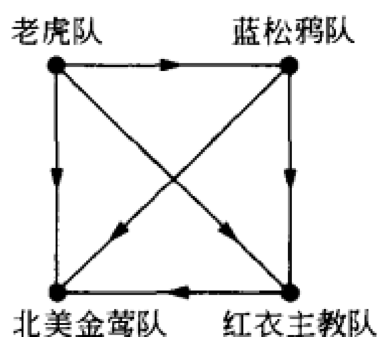
15.



17.



19.

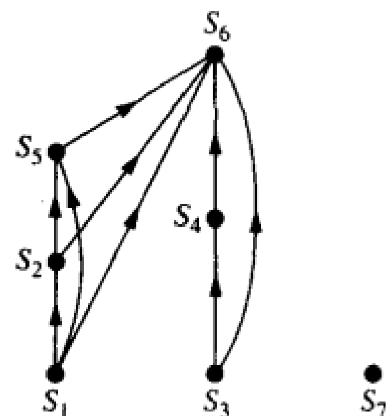


21. 找出在 2 月份的电话呼叫图中有但在 1 月份的电话呼叫图中却没有的电话号码，也找出在 1 月份有但在 2 月份却没有的号码。对于每个找出的号码，利用呼叫图中的边，列出每个号码的呼叫号码表或被叫号码表。检查这些表，以找出在与 1 月份停机的号码具有类似呼叫模式的 2 月份的新号码。

23. 利用这样的图模型，就是用电子邮件地址作为顶点，从发送地址到接收地址连一条边。对于每个电子邮件地址，可以建立其发送地址和接收地址表。如果两个电子邮件地址具有几乎相同的模式，就可以得出结论：这些地址可能属于同一个人，而该人最近改变了他或她的电子邮件地址。

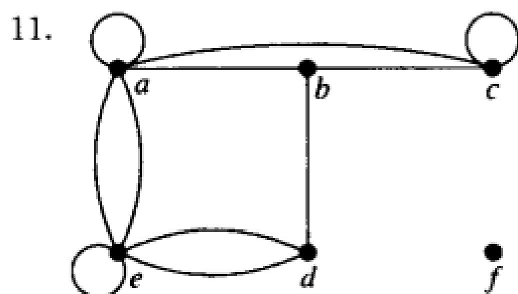
25. 令顶点集合为人的集合，如果两个人结过婚，则把两个顶点用边连起来。忽略掉同性婚姻，这个图就具有这样的性质，即有两种类型的顶点(男人和女人)，每条边都连接相反类型的顶点。

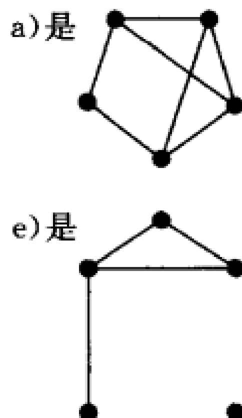
27.



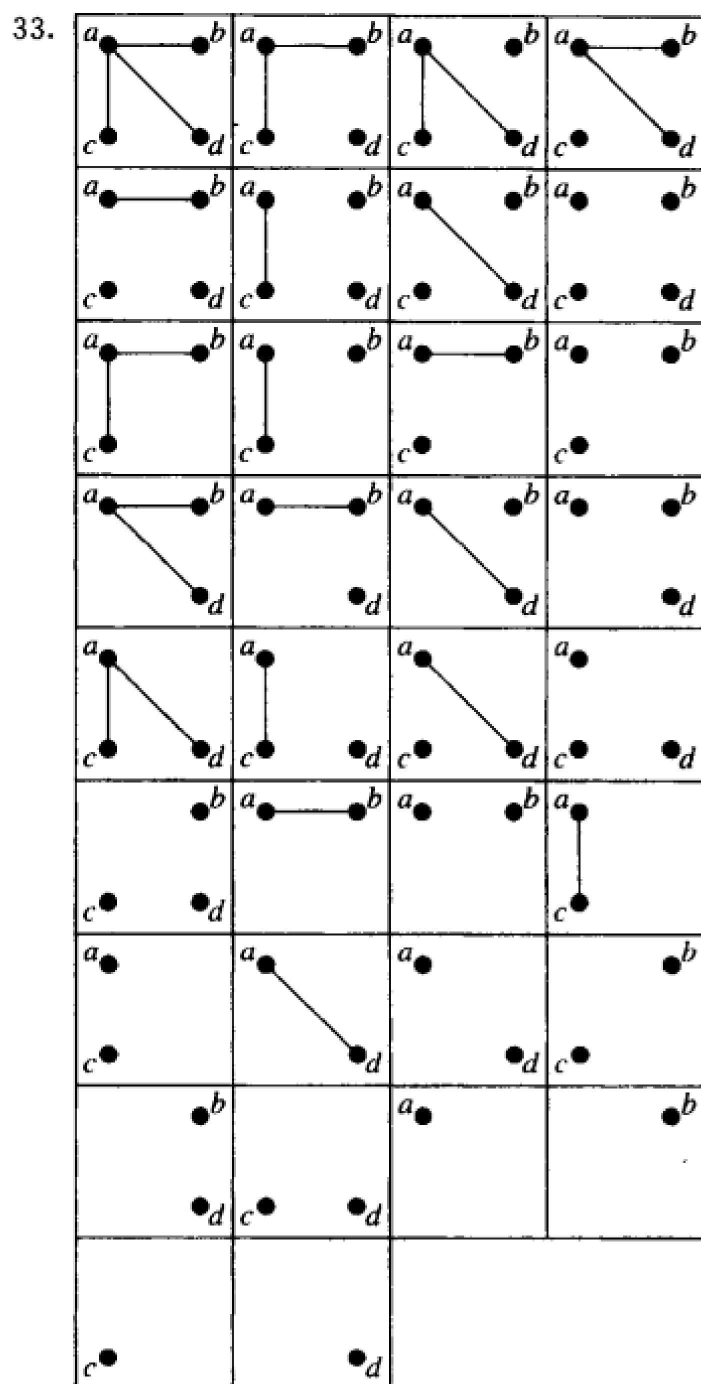
29. 用顶点表示组里的人们。对每对顶点来说，都在图里放置一条有向边。对从表示 A 的顶点到表示 B 的顶点的边来说，若 A 喜欢 B，则用＋(加)标记，若 A 不喜欢 B，则用－(减)标记，若 A 对 B 持中立态度，则用 0 标记。

8.2 节

1. $v=6$；$e=6$；$\deg(a)=2$，$\deg(b)=4$，$\deg(c)=1$，$\deg(d)=0$，$\deg(e)=2$，$\deg(f)=3$；c 是悬挂点；d 是孤立点。
3. $v=9$；$e=12$；$\deg(a)=3$，$\deg(b)=2$，$\deg(c)=4$，$\deg(d)=0$，$\deg(e)=6$，$\deg(f)=0$；$\deg(g)=4$；$\deg(h)=2$；$\deg(i)=3$；d 和 f 都是孤立点。
5. 否，因为顶点的度之和不可能是奇数。
7. $v=4$；$e=7$；$\deg^-(a)=3$，$\deg^-(b)=1$，$\deg^-(c)=2$，$\deg(d)=1$，$\deg^+(a)=1$，$\deg^+(b)=2$，$\deg^+(c)=1$，$\deg^+(d)=3$。
9. 5 个顶点，13 条边；$\deg^-(a)=6$，$\deg^+(a)=1$，$\deg^-(b)=1$，$\deg^+(b)=5$，$\deg^-(c)=2$，$\deg^+(c)=5$，$\deg^-(d)=4$，$\deg^+(d)=2$，$\deg^-(e)=0$，$\deg^+(e)=0$。
11. 

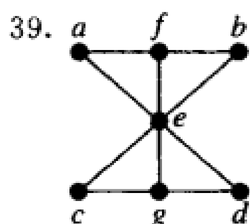
13. v 具有的合作者人数；从来没有合作过的某个人；只有一个合作者的某个人。
15. 在有向图中 $\deg^-(v)=v$ 收到的呼叫次数，$\deg^+(v)=v$ 发出的呼叫次数；在无向图中 $\deg(v)=$是 v 发出或收到的呼叫次数。
17. $(\deg^+(v), \deg^-(v))$是 v 的胜负记录。
19. 偶图　21. 非偶图　23. 非偶图
25. a)n 个顶点，$n(n-1)/2$ 条边　b)n 个顶点，n 条边
　c)$n+1$ 个顶点，$2n$ 条边　d)$m+n$ 个顶点，mn 条边
　e)2^n 个顶点，$n2^{n-1}$ 条边
27. 7
29. a)是　b)否，度之和是奇数。　c)否　d)否，度之和是奇数。
　e)是 　f)否，度之和是奇数。
31. 17

33.



35. a)对所有 $n \geqslant 1$　　b)对所有 $n \geqslant 3$　　c)对 $n=3$　　d)对所有 $n \geqslant 0$

37. 5　　39.



41. a)带 n 个顶点而且没有边的图。　　b) K_m 和 K_n 的不相交并图。

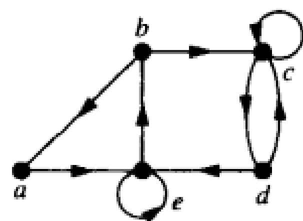
c)带有顶点 $\{v_1, \cdots, v_n\}$ 而且在 v_i 与 v_j 之间有边(除非 $i \equiv j \pm 1 (\bmod n)$)的图。

d)用长度为 n 的位串表示其顶点，若两个顶点所对应的位串相差超过一位，则在这两个顶点之间有一条边的图。

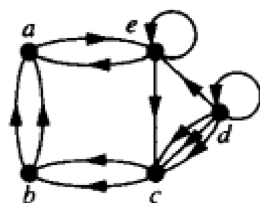
43. $v(v-1)/2-e$

45. G 和 $\bar{G}$ 的并图包含 n 个顶点中每对顶点之间的一条边。因此这个并图是 K_n。

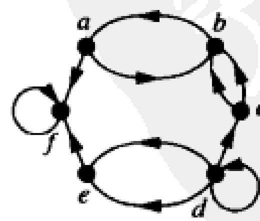
47. 练习 7：



练习 8：

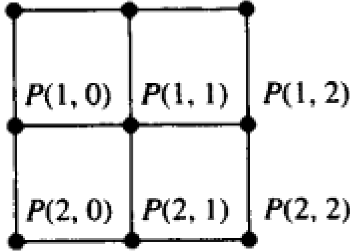


练习 9：



49. 有向图 $G=(V, E)$ 是它自身的逆图当且仅当它满足条件：$(u, v)\in E$ 当且仅当 $(v, u)\in E$。但这恰好是这样的条件：所对应的关系必定是对称的。

51.



53. 可以这样连接 $P(i, j)$ 与 $P(k, l)$：用 $|i-k|$ 个 hop 连接 $P(i, j)$ 与 $P(k, j)$，用 $|j-l|$ 个 hop 连接 $P(k, j)$ 与 $P(k, l)$。因此连接 $P(i, j)$ 与 $P(k, l)$ 所需要的 hop 总数不超过 $|i-k|+|j-l|$。这个值小于或等于 $m+m=2m$，它是 $O(m)$。

8.3 节

1.

顶点	相邻顶点
a	b，c，d
b	a，d
c	a，d
d	a，b，c

3.

顶点	相邻顶点
a	a，b，c，d
b	d
c	a，b
d	b，c，d

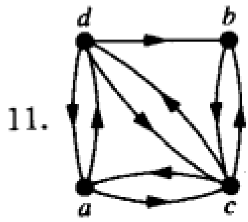
5. $\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$，其中顶点都以字母表顺序排列

7. $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \end{bmatrix}$

9. a) $\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$ b) $\begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$

c) $\begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}$ d) $\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$ e) $\begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix}$ f) $\begin{bmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$

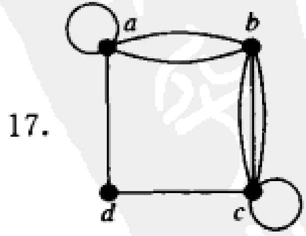
11.

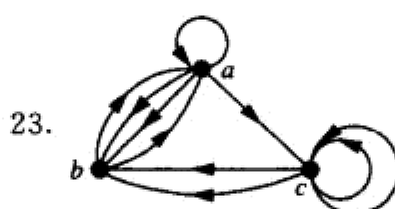


13. $\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 \\ 1 & 1 & 0 & 1 \\ 0 & 2 & 1 & 0 \end{bmatrix}$

15. $\begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 2 \\ 2 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{bmatrix}$

17.



19. $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$　　21. $\begin{bmatrix} 1 & 1 & 2 & 1 \\ 1 & 0 & 0 & 2 \\ 1 & 0 & 1 & 1 \\ 0 & 2 & 1 & 0 \end{bmatrix}$　　23. 

　　25. 是

27. 练习 13：$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$　　练习 14：$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$

练习 15：$\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \end{bmatrix}$

29. $\deg(v)$ − 在 v 处的环数；$\deg^-(v)$

31. 若 e 不是环则是 2，若 e 是环则是 1。

33. a) $\begin{bmatrix} 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 1 \end{bmatrix}$　　b) $\begin{bmatrix} 1 & 0 & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 1 & 1 \end{bmatrix}$

c) $\begin{bmatrix} 0 & 0 & \cdots & 0 & 1 & 1 & \cdots & 1 \\ & & & & 1 & 0 & \cdots & 0 \\ & \mathbf{B} & & & 0 & 1 & \cdots & 0 \\ & & & & \vdots & \vdots & & \vdots \\ & & & & 0 & 0 & \cdots & 1 \end{bmatrix}$，其中 **B** 是(b)的答案　　d) $\begin{bmatrix} 1 & 1 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 1 \\ 1 & 0 & & 0 & 1 & \cdots & 0 \\ 0 & 1 & & 0 & 0 & & \\ \vdots & \vdots & & \vdots & \vdots & & \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 1 \end{bmatrix}$

35. 同构　　37. 同构　　39. 同构　　41. 非同构　　43. 同构

45. G 通过恒等函数与自身同构，所以同构是自反的。假设 G 同构于 H，则存在从 G 到 H 的一一对应 f，f 保持相邻性与非相邻性。所以 f^{-1} 是从 H 到 G 的一一对应，f^{-1} 保持相邻性与非相邻性。因此同构是对称的。若 G 同构于 H 并且 H 同构于 K，则存在从 G 到 H 和从 H 到 K 的一一对应 f 和 g，f 和 g 保持相邻性与非相邻性。所以 $g \circ f$ 是从 G 到 K 的一一对应 f，$g \circ f$ 保持相邻性与非相邻性。因此同构是传递的。

47. 全是 0

49. 这样标记顶点，使得顶点集划分中第一个集合的所有顶点都排在前面。由于没有边连接同一划分集内的顶点，所有矩阵具有所需形式。

51. C_5　　53. 只有 $n=5$　　55. 4 个

57. a)同构　　b)非同构　　c)非同构

59. $G=(V_1, E_1)$ 同构于 $H=(V_2, E_2)$，当且仅当存在从 V_1 到 V_2 的函数 f 和从 E_1 到 E_2 的函数 g，使得 f 和 g 都是一一对应，并且对于 E_1 中的每条边 e，$g(e)$ 的端点是 $f(v)$ 和 $f(w)$，其中 v 和 w 是 e 的端点。

61. 同构　　63. 同构

65. 若 f 是从有向图 G 到有向图 H 的同构，则 f 也是从 G^c 到 H^c 的同构。为了看出这一点，注意 (u, v) 是 G^c 的边，当且仅当 (u, v) 是 G 的边，当且仅当 $(f(u), f(v))$ 是 H 的边，当且仅当 $(f(u), f(v))$ 是 H^c 的边。

67. 乘积矩阵是$[a_{ij}]$，其中当$i\neq j$时，a_{ij}是从v_i到v_j的边数，a_{ii}是与v_i关联的边数。
69. 练习41中的图就是魔鬼对。

8.4 节

1. a)长度为4的通路；非回路；非简单通路　b)非通路
c)非通路　d)长度为5的简单回路
3. 非连通　5. 非连通
7. 具有下列性质的人的极大集：对于其中任意二人，可以找到从一个人通向另一个人的熟人关系串。
9. 如果一个人的爱尔特希数为n，则在合作关系图中存在从该人到爱尔特希的长为n的路径，根据定义，这意味着该人与爱尔特希在同一个分支中。如果一个人与爱尔特希在同一个分支中，则存在从该人到爱尔特希的路径，最短的这种路径的长度就是该人的爱尔特希数。
11. 电话号码的极大集，对于其中任意两个不同的号码，可以找出二者之间的有向路径。
13. 求下列每个图的强连通分支。
a). $\{a, b, f\}$，$\{c, d, e\}$　b). $\{a, b, c, d, e, h\}$，$\{f\}$，$\{g\}$　c). $\{a, b, d, e, f, g, h, i\}$，$\{c\}$
15. a)2　b)7　c)20　d)61　17. a)3　b)0　c)27　d)0
19. a)1　b)0　c)2　d)1　e)5　f)3
21. 由定义，R是自反的。假设$(u, v)\in R$，则存在从u到v的通路。于是$(v, u)\in R$，因为存在从v到u的通路，即从u到v的通路的倒转。所以R是对称的。假设$(u, v)\in R$且$(v, w)\in R$，则存在从u到v和从v到w的通路。把这两条通路合在一起就给出从u到w的通路，因此$(u, w)\in R$，所以R是传递的。
23. c　25. b，c，e，i
27. 如果一个顶点是悬挂点，显然它不是割点。所以割边的端点是割点就不是悬挂点。删除一条割边，产生比原图有更多连通分支的图，如果割边的端点不是悬挂点，则在删除割边后，该端点所在的连通分支就不只含有这一个顶点。因此，删除该顶点及其关联的所有边，包括原来那条割边，就产生比原图有更多连通分支的图。因此，割边的端点不是悬挂点就是割点。
29. 假设连通图G至多有1个顶点不是割点。定义顶点u和v之间的距离（记作$d(u, v)$）是G中u与v之间最短通路的长度。设s和t是G中使得$d(s, t)$最大的顶点。s或t之一（或二者都）是割点。所以不失一般性，假设s是割点。在从G删除s及其关联边所得到的图中，设w属于不含t的那个连通分支。由于从w到t的每条通路都含有s，所以$d(w, t)>d(s, t)$，这是矛盾。
31. a)丹佛-芝加哥，波士顿-纽约
b)西雅图-波特兰，波特兰-旧金山，盐湖城-丹佛，纽约-波士顿，波士顿-伯灵顿，波士顿-班戈
33. 一组人合起来能（直接或间接）影响每个人；{Deborah}
35. 一条边不能连接不同连通分支内的两个顶点。由于在有n_i个顶点的连通分支内至多有$C(n_i, 2)$条边，所以图中至多有$\sum_{i=1}^{k} C(n_i, 2)$条边。
37. 假设G不连通。则有一个连通分支有k个顶点，$1\leqslant k\leqslant n-1$。$G$的边数最多为$C(k, 2)+C(n-k, 2)=(k(k-1)+(n-k)(n-k-1))/2=k^2-nk+(n^2-n)/2$。这个$k$的二次函数在$k=n/2$处最小，在$k=1$或$k=n-1$处最大。因此，若$G$不连通，则边数不超过这个函数在1和$n-1$处的值，即$(n-1)(n-2)/2$。
39. a)1　b)2　c)6　d)21
41. 2
43. 设通路P_1和P_2分别是$u=x_0, x_1, \cdots, x_n=v$和$u=y_0, y_1, \cdots, y_m=v$。由于$P_1$和$P_2$不含公共边，所以迟早会分开。如果在其中一条通路结束后才分开，则另一条通路的剩余部分就是从v到u的回路。否则，假设$x_0=y_0, x_1=y_1, \cdots, x_i=y_i$，但是$x_{i+1}\neq y_{i+1}$。沿着通路$y_i, y_{i+1}, y_{i+2}, \cdots$前进，直到再次遇到$P_1$上的顶点为止。一旦回到$P_1$，就沿着$P_1$向前或向后（视需要而定），直到回到$x_i$。由于

$x_i = y_i$，这样形成的回路必定是简单的，因为诸 x_k 之间的边没有重复，而且这些边也不同于所用到的诸 y_i 之间的边。

45. 图是连通的当且仅当 $\mathbf{A}+\mathbf{A}^2+\mathbf{A}^3+\cdots+\mathbf{A}^{n-1}$ 的对角线外元素都是正的，其中 $\mathbf{A}$ 是 G 的邻接矩阵。

47. 如果图是偶图，不妨设顶点分成 A 和 B 两部分，则每条通路上的顶点必定交替属于 A 和 B。因此，从比方说 A 出发的通路在奇数步后停在 B，在偶数步后停在 A。由于回路总是停在出发的同一个顶点上，所以长度必为偶数。反之，假设所有回路长度都是偶数，要证图是偶图。可以假设图是连通的，因为若不连通，则每次仅考虑一个连通分支即可。设 v 是图的一个顶点，设 A 是有从 v 出发奇数长度通路的所有顶点的集合，设 B 是有从 v 出发偶数长度通路的所有顶点的集合。由于这个分支是连通的，所以每个顶点都属于 A 或 B。没有顶点同时属于 A 和 B，因为那样一来，从 v 到那个顶点的奇长度通路，加上从那个顶点到 v 的偶长度通路，就产生一个奇回路，与假设矛盾。因此，顶点集合划分成两部分。为了证明每条边的端点都在不同部分中，假设 xy 是一条边，其中 $x\in A$。则从 v 到 x 的奇长度通路加上 xy，就产生从 v 到 y 的偶长度通路，所以 $y\in B$。(若 $x\in B$，也类似可证。)

8.5 节

1. 均无
3. 无欧拉回路；$a, e, c, e, b, e, d, b, a, c, d$
5. $a, b, c, d, c, e, d, b, e, a, e, a$
7. $a, i, h, g, d, e, f, g, c, e, h, d, c, a, b, i, c, b, h, a$
9. 不能，A 仍然是奇数度的。
11. 当顶点表示交叉路口，边表示街道的图有欧拉通路时。
13. 能一笔画　　15. 不能一笔画
17. 如果有欧拉通路，则沿着该通路前进时，除起点和终点外的每个顶点都有相同的入度和出度，因为每当沿一条边进入一个顶点时，还要沿另一条边离开它。起点必定是出度比入度大 1，因为用一条边离开这个顶点，而每当再次访问它时，要用一条边进入和一条边离开。同理，终点必定是入度比出度大 1。由于不考虑方向的欧拉通路在无向底图中产生任意两个顶点之间的通路，所以该图是弱连通的。反之，假设图满足题设条件。如果从出度小的顶点到入度小的顶点之间增加一条边，则该图每个顶点入度与出度相等。由于该图仍然是弱连通的，根据练习 16，这个新图有欧拉回路。现在删除后加的边，就得到欧拉通路。
19. 均无
21. 无欧拉回路；$a, d, e, d, b, a, e, c, e, b, c, b, e$
23. 均无
25. 与算法 1 的过程基本相同，要考虑遵循边的方向。
27. a) $n=2$　　b) 无　　c) 无　　d) $n=1$
29. 练习 1：1 次；练习 2～7：0 次
31. a, b, c, d, e, a 是哈密顿回路
33. 不存在哈密顿回路，因为一旦这样的回路到达 e，就无路可走了。
35. 不存在哈密顿回路，因为图中每条边都与 2 度点关联，因此必须在该回路上。
37. a, b, c, f, d, e 是哈密顿通路
39. f, e, d, a, b, c 是哈密顿通路
41. 不存在哈密顿通路。有八个 2 度点，其中只有两个是通路的端点。对另外六点中的每个点，两条关联边必定在通路上。不难看出，如果有哈密顿通路，则恰好有一个角落里的顶点是端点，但这是不可能的。
43. $a, b, c, f, i, h, g, d, e$ 是哈密顿通路。
45. $m=n\geqslant 2$

47. a)不能；不能；无　b)不能；不能；有　c)能；能；有　d)能；能；有

49. 对于 $n=1$，结果是平凡的，编码是 0，1。假设存在 n 阶格雷码，设 c_1，c_2，…，c_k 是这样的编码，$k=2^n$。则 $0c_1$，$0c_2$，…，$0c_k$，$1c_1$，$1c_2$，…，$1c_k$ 是 $n+1$ 阶格雷码。

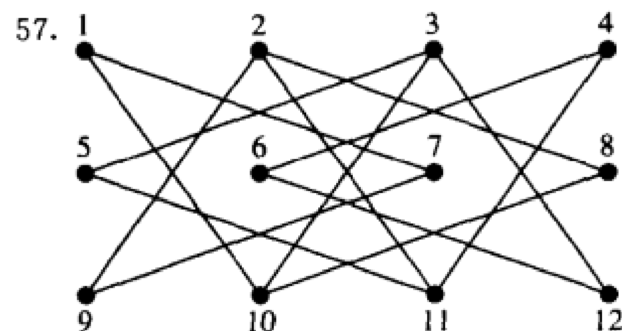
51. **procedure** *Fleury*($G=(V, E)$：连通多重图，所有顶点度为偶数，$V=\{v_1, \cdots, v_n\}$)

```
v := v1
circuit := v
H := G
while H 有边
begin
  e := 若 H 中存在端点为 v 且不是 H 中割边的边，则是第一条这样的边(按照 V 的排序)，否则就是
  H 中端点为 v 的第一条边
   w := e 的另一个端点
   circuit := circuit 加入边 e 和顶点 w
   v := w
   H := H - e
end{circuit 是欧拉回路}
```

53. 若 G 有欧拉回路，则该回路也是欧拉通路。若无欧拉回路，则在两个奇数度顶点之间加一条边，应用算法求出欧拉回路。然后删除新加的边。

55. 假设 $G=(V, E)$ 是满足 $V=V_1\cup V_2$ 的偶图，其中没有边是连接 V_1 里的顶点与 V_2 里的顶点的。假设 G 有哈密顿回路。这样的回路必然是形如 a_1，b_1，a_2，b_2，…，a_k，b_k，a_1，b_1，其中对 $i=1, 2, \cdots, k$ 来说，有 $a_i\in V_1$ 和 $b_i\in V_2$。因为哈密顿回路访问每个顶点恰好一次，除了回路开始和结束的 a_1 之外，所以图中的顶点数等于 $2k$，它是偶数。因此，带奇数个顶点的偶图不可能有哈密顿回路。

57.



59. 把 3×4 棋盘的格子表示如下：

1	2	3	4
5	6	7	8
9	10	11	12

马的巡回路线可以通过下列移动来构成：8，10，1，7，9，2，11，5，3，12，6，4。

61. 把 4×4 棋盘的格子表示如下：

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

从角上的四个格子里只有两种移动。若包含 1-10，1-7，16-10 和 16-7 的所有边，则过早地完成了回路，所以必须去掉这些边中的至少一条。不失一般性，假设通路开始于 1-10，10-16 和 16-7。现在从格子 3 仅有的移动是到格子 5，10 和 12，而格子 10 已经有了两条关联的边。因此 3-5 和 3-12 必须是在哈密顿回路里。同理，边 8-2 和 8-15 必须是在回路里。现在从格子 9 仅有的移动是到格子 2，7 和 15。假如有从格子 9 到格子 2 和 15 的边，则过早地完成了回路。因此边 9-7 必须是在回路里，让格子 7 关联的边达到饱和。但是现在格子 14 被迫连接到格子 5 和 12，过早地完成了回路(5-14-12-3-5)。这个矛盾证明在 4×4 棋盘上没有马的巡回路线。

63. 因为在 $m\times n$ 棋盘上有 mn 个格子，所以若 m 和 n 都是奇数，则有奇数个格子。因为根据练习 62，对应的图是偶图，所以根据练习 55，它没有哈密顿回路。因此没有可重入的马的巡回路线。

65. a)如果 G 没有哈密顿回路，则不断地尽量把 G 没有的边逐条加入 G，条件是产生的图不含哈密顿回路。不可能一直这样做下去，因为一旦把 G 没有的所有边都加入了，就形成了完全图，就有了哈密顿回路。到这个过程停止时，就得到了具有所需要性质的图 H(肯定是非完全的)。

 b)再给 H 加入一条边。这样就产生了经过所加入的边的哈密顿回路。从这个回路上去掉加入的边所形成的通路，就是 H 中的哈密顿通路。

 c)显然 v_1 与 v_n 在 H 中不相邻，因为 H 没有哈密顿回路。因此 v_1 与 v_n 在 G 中不相邻。但是假设了在 G 中不相邻顶点的度之和至少为 n。这个不等式可以改写成 $n-\deg(v_n)\leqslant\deg(v_1)$。但 $n-\deg(v_n)$ 正是不与 v_n 相邻的顶点数。

 d)由于在哈密顿通路上 v_n 没有后续顶点，所以 v_n 不属于 S。与 v_1 相邻的 $\deg(v_1)$ 个顶点中每个顶点都引出一个 S 元素，所以 S 包含 $\deg(v_1)$ 个顶点。

 e)根据(c)部分，至多有 $\deg(v_1)-1$ 个与 v_n 不同且不与 v_n 相邻的顶点，根据(d)部分，S 中有 $\deg(v_1)$ 个顶点，这些顶点都不是 v_n。因此至少有一个 S 中的顶点与 v_n 相邻。根据定义，如果 v_k 是这个顶点，则 H 包含边 v_kv_n 和 v_1v_{k+1}，其中 $1<k<n-1$。

 f)现在 $v_1, v_2, \cdots, v_{k-1}, v_k, v_n, v_{n-1}, \cdots, v_{k+1}, v_1$ 是 H 中的哈密顿回路，与 H 的构造矛盾。因此 G 原来没有哈密顿回路的假设是错误的，反证法完成了。

8.6 节

1. a)顶点是车站，边连接相邻车站，权是相邻车站之间旅行所需要的时间。
 b)与(a)基本相同，不同之处在于，权是相邻车站之间的距离。
 c)与(a)基本相同，不同之处在于，权是车站之间的票价。

3. 16

5. 练习 2：a, b, e, d, z；练习 3：a, c, d, e, g, z；练习 4：$a, b, e, h, l, m, p, s, z$

7. a)a, c, d b)a, c, d, f c)c, d, f d)b, d, e, g, z

9. a)直飞 b)经纽约 c)经亚特兰大和芝加哥 d)经纽约

11. a)经芝加哥 b)经芝加哥 c)经洛杉矶 d)经芝加哥

13. a)经芝加哥 b)经芝加哥 c)经洛杉矶 d)经芝加哥。

15. 当把 z 加入集合 S 时，算法不停止。

17. a)经木桥，经木桥和康登 b)经木桥，经木桥和康登

19. 例如，观光路线，清扫街道

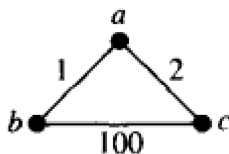
21.

	a	*b*	*c*	*d*	*e*	*z*
a	4	3	2	8	10	13
b	3	2	1	5	7	10
c	2	1	2	6	8	11
d	8	5	6	4	2	5
e	10	7	8	2	4	3
z	13	10	11	5	3	6

23. $O(n^3)$

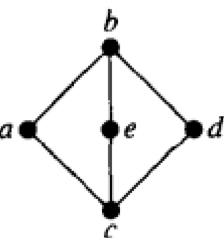
25. a-c-b-d-a(或者从某个点开始，以相同或相反顺序遍历这些顶点的相同回路)

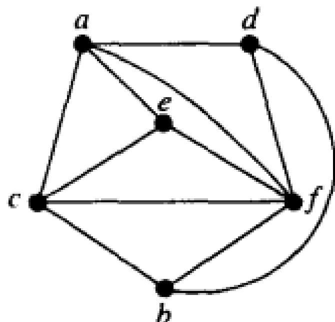
27. 旧金山–丹佛–底特律–纽约–洛杉矶–旧金山(或者从某个点开始，以相同或相反顺序遍历这些顶点的相同回路)

29. 考虑这个图：

回路 a-b-a-c-a 访问每个顶点至少一次(访问顶点 a 两次)，总权为 6。每个哈密顿回路有总权 103。

8.7 节

1. 能　3. 

5. 非可平面的　7. 可平面的 

9. 非可平面的

11. 由连接 v_1，v_2，v_3 的边组成的 K_5 子图，其平面表示形成一个三角形。顶点 v_4 要么在三角形之内，要么在三角形之外。下面只考虑 v_4 在三角形内的情形，另一种情形是类似的。画出从 v_1，v_2，v_3 到 v_4 的边，就一共形成 4 个区域。无论 v_5 在哪个区域中，都只能让 v_5 与其余顶点中的 3 个顶点相连，而不能与所有 4 个顶点相连。

13. 8

15. 由于没有环和多重边，也没有长度为 3 的回路，而且无界区域的次数至少为 4，所以每个区域的次数至少为 4。因此，$2e \geqslant 4r$，或 $r \leqslant e/2$。但是 $r = e - v + 2$，所以有 $e - v + 2 \leqslant e/2$，这蕴含着 $e \leqslant 2v - 4$。

17. 与推论 2 中一样，有 $2e \geqslant 5r$ 和 $r = e - v + 2$，因此 $e - v + 2 \leqslant 2e/5$，这蕴含着 $e \leqslant (5/3)v - (10/3)$。

19. 只有 a)和 c)　21. 不同胚于 $K_{3,3}$　23. 可平面的　25. 非可平面的

27. a)1　b)3　c)9　d)2　e)4　f)16

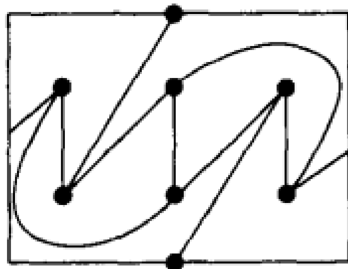
29. 按照提示来画 $K_{m,n}$。交叉数是第一象限中数目的 4 倍。在 x 轴上原点右边的顶点是(1, 0)，(2, 0)，…，$(m/2, 0)$，在 y 轴上原点上边的顶点是(0, 1)，(0, 2)，…，$(0, n/2)$。选择任意两个不同的数 a 和 b，$1 \leqslant a < b \leqslant m/2$，以及两个不同的数 r 和 s，$1 \leqslant r < s \leqslant n/2$，就得出所有的交叉。在图中连接$(a, 0)$和$(0, s)$的边与连接$(b, 0)$和$(0, r)$的边之间，恰好产生一处交叉。因此，在第一象限中的交叉数是 $C\left(\frac{m}{2}, 2\right) \cdot C\left(\frac{n}{2}, 2\right) = \frac{(m/2)(m/2-1)}{2} \cdot \frac{(n/2)(n/2-1)}{2}$。因此，总交叉数是 $4 \cdot mn(m-2)(n-2)/64 = mn(m-2)(n-2)/16$。

31. a)2　b)2　c)2　d)2　e)2　f)2

33. 对于 $n \leqslant 4$，公式是正确的。若 $n > 4$，根据练习 32，K_n 的厚度至少是 $C(n, 2)/(3n-6) = \left(n+1+\frac{2}{n-2}\right)\Big/6$ 的上取整。由于 $\left(n+1+\frac{2}{n-2}\right)\Big/6$ 永远不是整数，它就等于加 1 后下取整，即等于 $\lfloor (n+7)/6 \rfloor$。

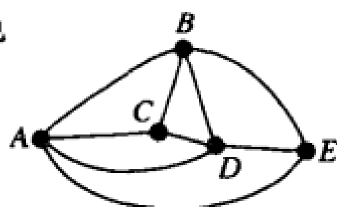
35. 这个结果来自练习 34，因为 $K_{m,n}$ 有 mn 条边和 $m+n$ 个顶点，并且由于是偶图而没有三角形。

37.

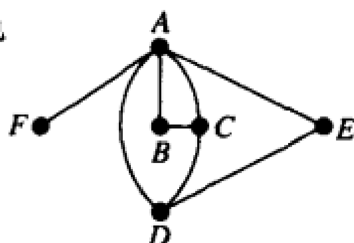


8.8 节

1. 4 色



3. 3 色



5. 3　7. 3　9. 2　11. 3

13. 没有边的图　15. 3(若 n 为偶数)，4(若 n 为奇数)

17. 时间段 1：Math 115，Math 185；时间段 2：Math 116，CS 473；时间段 3：Math195，CS 101；时间段 4：CS 102；时间段 5：CS 273

19. 5

21. 练习 5：3，练习 6：6，练习 7：3，练习 8：4，练习 9：3，练习 10：6，练习 11：4

23. 5

25. 具有一种颜色的顶点集是一部分，具有另一种颜色的顶点集是另一部分。由于没有边连接相同颜色的顶点，所以在同一部分的顶点之间没有边。

27. 颜色 1：e，f，d；颜色 2：c，a，i，g；颜色 3：h，b，i

29. 给 C_6 着色　31. a) 6　b) 7　c) 9　d) 11

33. 用颜色表示频率，用顶点表示地段。如果两个顶点表示的地段互相干涉，则用边连接这两个顶点。于是 k 重着色恰好就是避免了干涉的频率分配。

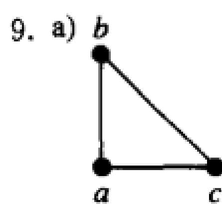
35. 对图的顶点数进行归纳。具有 5 个或少于 5 个顶点的图可用 5 色或少于 5 色来着色，因为每个顶点可用不同的颜色。这就完成了基础情形。假设所有 k 个顶点的图都可 5 着色，考虑$k+1$个顶点的图 G。根据 8.7 节推论 2，G 至少有一个度不超过 5 的顶点 v。删除 v 得到图 G'。由于 G' 只有 k 个顶点，由归纳假设可对 G' 做 5 着色。如果与 v 相邻的顶点没有用完 5 种颜色，则把 v 的邻点所没有用的颜色给 v，就给 G 做了 5 着色。麻烦出现在若 v 有 5 个邻点，并且每个点用了 G' 的 5 着色中的一种不同颜色。假设以逆时针顺序考虑 v 的邻点是 a，b，c，m，p。(表示与 v 关联的边的曲线的逆时针顺序决定了这个顶点顺序。)假设这些邻点的颜色分别是：天蓝(azure)，蓝(blue)，黄绿(chartreuse)，洋红(magenta)，紫(purple)。考虑天蓝-黄绿子图(即 G 中着天蓝或黄绿色的顶点以及这些顶点之间的边)。如果 a 和 c 在这个子图的不同连同分支中，则在含有 a 的分支中可以交换这两种颜色(让天蓝色顶点改成黄绿色，让黄绿色顶点改成天蓝色)，而 G' 仍然是适当着色的。这样就让 a 成为黄绿色，于是现在可给 v 着天蓝色，G 就适当着色了。如果 a 和 c 在同一个连通分支中，则有一条路径连接 a 和 c，其上的顶点交替着天蓝和黄绿色。这条路径连同边 av 和 vc 就把平面划分成两个区域，其中一个区域含有 b，另外一个区域含有 m。如果现在给 b 所在区域中的所有顶点交换蓝色和洋红色，则仍然得到 G' 的恰当着色，但是现在蓝色可给 v 着色了。在这种情况下，也找到了 G 的恰当着色。这样就完成了归纳步骤，于是定理得证。

补充练习

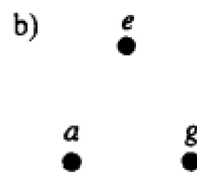
1. 2500　3. 同构　5. 同构

7. $\sum_{i=1}^{m} n_i$ 个顶点，$\sum_{i<j} n_i n_j$ 条边

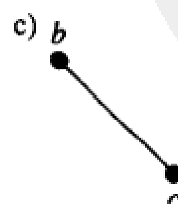
9. a)



b)



c)



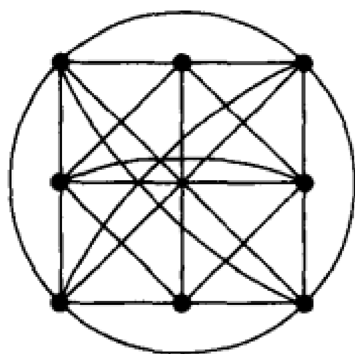


11. 完全子图含有下列顶点集合：$\{b, c, e, f\}$，$\{a, b, g\}$，$\{a, d, g\}$，$\{d, e, g\}$，$\{b, e, g\}$

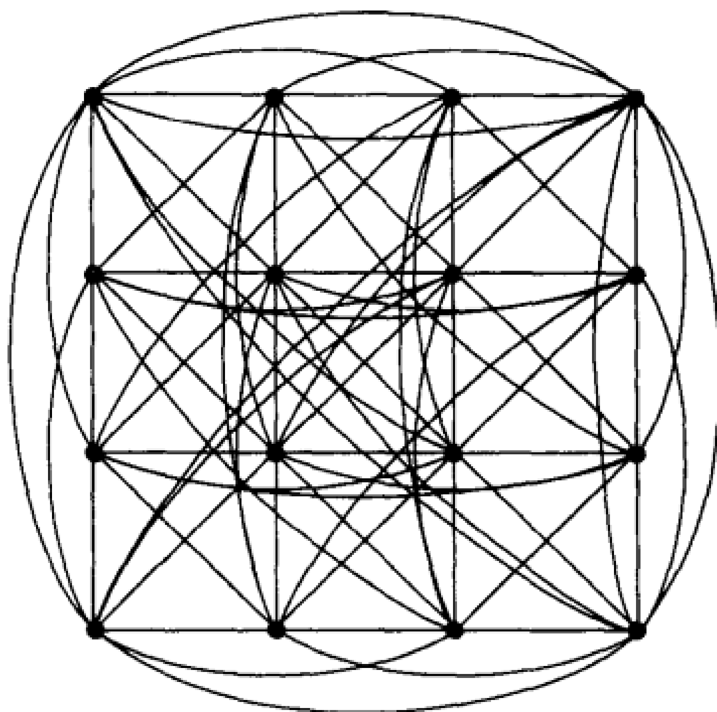
13. 完全子图含有下列顶点集合：$\{b, c, d, j, k\}$，$\{a, b, j, k\}$，$\{e, f, g, i\}$，$\{a, b, i\}$，$\{a, i, j\}$，$\{b, d, e\}$，$\{b, e, i\}$，$\{b, i, j\}$，$\{g, h, i\}$，$\{h, i, j\}$

15. $\{c, d\}$是最小支配集

17. a)



b)



19. a)1　　b)2　　c)3

21. a)图 G 中从 u 到 v 的路径导出同构图 H 中从 $f(u)$ 到 $f(v)$ 的路径。

b)假设 f 是从 G 到 H 的同构映射。若 v_0，v_1，…，v_n，v_0 是 G 中的哈密顿回路，则 $f(v_0)$，$f(v_1)$，…，$f(v_n)$，$f(v_0)$ 必为 H 中的哈密顿回路，因为它还是回路并且对于 $0 \leqslant i < j \leqslant n$，$f(v_i) \neq f(v_j)$。

c)假设 f 是从 G 到 H 的同构映射。若 v_0，v_1，…，v_n，v_0 是 G 中的欧拉回路，则 $f(v_0)$，$f(v_1)$，…，$f(v_n)$，$f(v_0)$ 必为 H 中的欧拉回路，因为它还是过每条边恰好一次的回路。

d)两个同构图必有相同的交叉数，因为它们可用完全相同的方式画在平面上。

e)假设 f 是从 G 到 H 的同构映射，则 v 在 G 中是孤立点当且仅当 $f(v)$ 在 H 中是孤立点。因此这两个图必有相同的孤立点数。

f)假设 f 是从 G 到 H 的同构映射。若 G 是偶图，则 G 的顶点集可划分成 V_1 和 V_2，使得没有边连接一个 V_1 中的顶点和一个 V_2 中的顶点。于是 H 的顶点集可划分成 $f(V_1)$ 和 $f(V_2)$，使得没有边连接一个 $f(V_1)$ 中的顶点和一个 $f(V_2)$ 中的顶点。

23. 3 种

25. a)自逆　　b)非自逆

27. 不可定向　　29. 可定向

31. 如果 e 是端点为 u 和 v 的割边，那么若 e 的方向是从 u 到 v，则在该有向图中没有从 v 到 u 的路径，否则 e 就不是割边。若 e 的方向是从 v 到 u，则可做类似的推理。

33. $n-1$

35. 让顶点表示鸡，图中有边 (u, v) 当且仅当鸡 u 比鸡 v 占优势。

37. a)4　　b)2　　c)3　　d)4　　e)4　　f)2

39. a)假设 $G=(V, E)$，设 a，$b \in V$。要证明 a 和 b 在 $\overline{G}$ 中距离至多为 2。若 $\{a, b\} \notin E$，则此距离为 1，所以假设 $\{a, b\} \in E$。由于 G 的直径大于 3，所以存在顶点 u 和 v，使得 u 和 v 之间在 G 中的距离大于 3。u 或 v 二者之一或全部都不属于集合 $\{a, b\}$。假设 u 与 a，b 不同，则 $\{a, u\}$ 或 $\{b, u\}$ 之一属于 E；否则 a，u，b 就是 $\overline{G}$ 中长度为 2 的路径。所以，不失一般性，假设 $\{a, u\} \in E$。因此 v 不是 a 或 b，而根据同样的推理，要么 $\{a, v\} \in E$，要么 $\{b, v\} \in E$。无论哪种情形，都给出 G 中从 u 到 v 长度小于等于 3 的路径，这是矛盾。

b)假设 $G=(V, E)$，设 a，$b \in V$。要证明 a 和 b 在 $\overline{G}$ 中距离不超过 3。若 $\{a, b\} \notin E$，则得出所要结果，所以假设 $\{a, b\} \in E$。由于 G 的直径大于等于 3，所以存在顶点 u 和 v，使得 u 和 v 之间在 G 中的距离大于等于 3。u 或 v 二者之一或全部都不属于集合 $\{a, b\}$。假设 u 与 a，b 不同。要么 $\{a, u\} \in E$，要么 $\{b, u\} \in E$；否则 a，u，b 就是 $\overline{G}$ 中长度为 2 的路径。所以，不失一般性，假设 $\{a, u\} \in E$。因此 v 与 a 和 b 都不同。若 $\{a, v\} \in E$，则 u，a，v 是 G 中长度为 2 的路径，所以 $\{a, v\} \notin E$，因此 $\{b, v\} \in E$(否则 $\overline{G}$ 中有长度为 2 的路径 a，v，b)。因此 $\{u, b\} \notin E$；否则 u，b，v 是 G 中长度为 2 的

路径。因此，a，v，u，b 是 $\overline{G}$ 中长度为 3 的路径，即为所求。

41. a，b，e，z　　43. a，c，b，d，e，z

45. 若 G 是可平面的，则由于 $e\leqslant 3v-6$，G 至多有 27 条边。(若 G 不连通，G 的边还要少。)同理，$\overline{G}$ 至多有 27 条边。但是 G 和 $\overline{G}$ 的并图是 K_{11}，K_{11} 有 55 条边，而 $55>27+27$。

47. 假设 G 用 k 色着色并且独立数为 i。由于每个颜色类都是独立集，所以每个颜色类有不超过 i 个元素。因此至多有 ki 个顶点。

49. a)根据 5.2 节定理 2，恰好选中 m 条边的概率是 $C(n, m)p^m(1-p)^{n-m}$。

b)根据 5.3 节定理 2，期望值是 np。

c)为了生成一个带标记图 G，当把这个过程应用到顶点对上时，当那对顶点之间有 G 边时，所选择的随机数 x 必定小于等于 1/2；当那里无 G 边时，x 大于 1/2。因此，做正确选择的概率对每条边来说是 1/2，对所有边来说是 $1/2^{C(n,2)}$。因此所有带标记的图是等概率的。

51. 假设 P 是单调增的。如果从简单图中删除边时，不保持不具有性质 P 这个性质，就有一个简单图 G 不具有性质 P，有另外一个简单图 G' 具有性质 P，G' 与 G 有相同顶点但是比 G 少一些边。但是性质 P 是单调增的，由于 G' 具有性质 P，所以给 G' 添加边得到的 G 也具有性质 P。这是矛盾。逆命题的证明是类似的。

第 9 章

9.1 节

1. (a)，(c)，(e)

3. a)a　　b)a，b，c，d，f，h，j，q，t　　c)e，g，i，k，l，m，n，o，p，r，s，u

d)q，r　　e)c　　f)p　　g)f，b，a　　h)e，f，l，m，n

5. 否

7. 0 层：a；1 层：b，c，d；2 层：e 到 k(按字母顺序)；3 层：l 到 r；4 层：s，t；5 层：u

9. a)整棵树　b)c，g，h，o，p 和 4 条边 cg，ch，ho，hp　c)只有 e

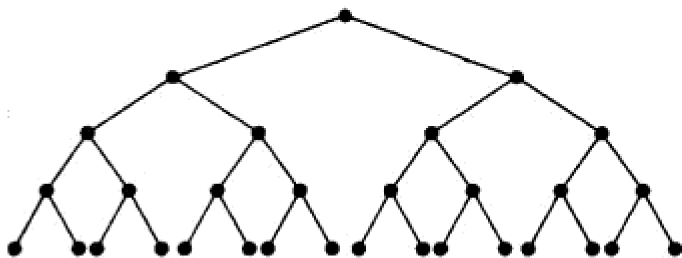
11. a)1　　b)2　　13. a)3　　b)9

15. "仅当"部分就是定理 2 和树的定义。假设 G 是连通简单图，有 n 个顶点和 $n-1$ 条边。如果 G 不是树，由练习 14，G 包含这样一条边，删除这条边产生一个图 G'，G' 仍连通。如果 G' 不是树，删除一条边产生连通图 G''。重复这个过程直到得出树为止。这至多需要 $n-1$ 步，因为只有 $n-1$ 条边。由定理 2，得出的图有 $n-1$ 条边，因为它有 n 个顶点。于是并没有删除边，所以 G 本身就是树。

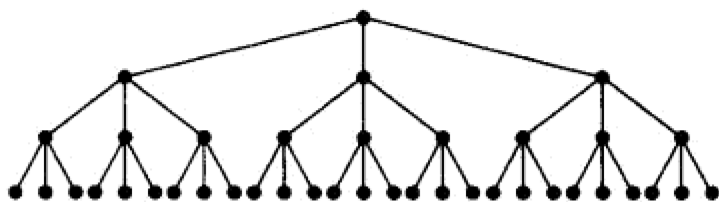
17. 9999　　19. 2000　　21. 999　　23. 1 000 000 美元

25. 由定理 4 不存在这样的树，因为 $m=2$ 或 $m=84$ 是不可能的。

27. 高度为 4 的完全二叉树



高度为 3 的完全 3 元树



29. a)根据定理 3，有 $n=mi+1$。因为 $i+l=n$，所以有 $l=n-i$，所以 $l=(mi+1)-i=(m-1)i+1$。

b)有 $n=mi+1$ 和 $i+l=n$，因此 $i=n-l$，$n=m(n-l)+1$。对 n 求解就给出 $n=(ml-1)(m-1)$。从 $i=n-l$ 得出 $i=[(ml-1)/(m-1)]-l=(l-1)/(m-1)$。

31. $n-t$　　33. a)1　　b)3　　c)5

35. a)父目录

b)子目录或包含的文件

c)在相同父目录里的子目录或包含的文件

d)在路径名称里的所有目录

e)在该目录或该目录的子目录里的所有子目录和包含的文件，依此类推

f)到这个目录或文件的路径的长度

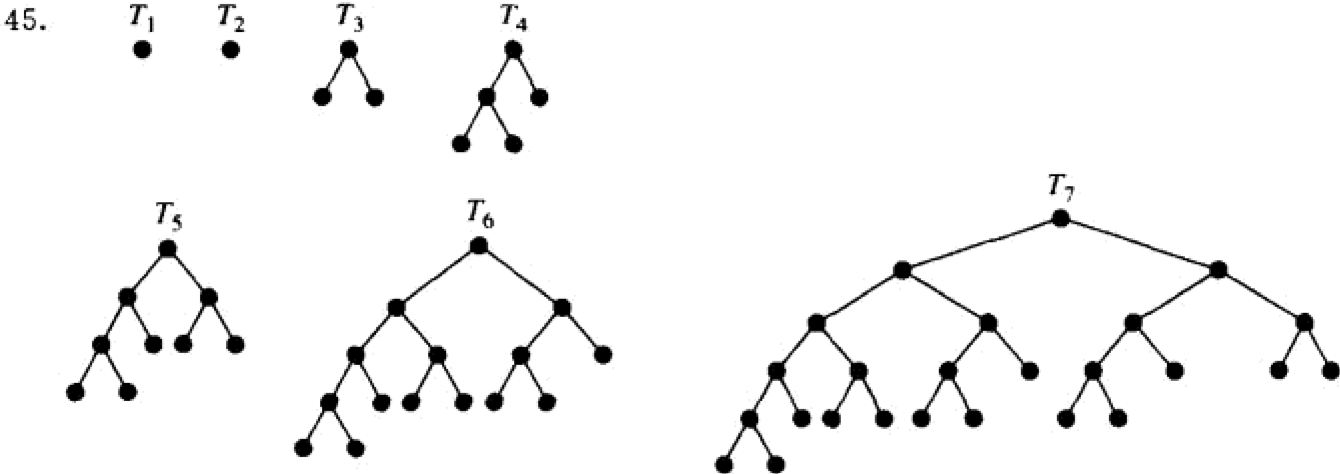
g)系统的深度，即最长路径的长度

37. 设 $n=2^k$，其中 k 是正整数。若 $k=1$，则没有什么要证明的，因为可以用 $n-1=1$ 个处理器在 $\log 2=1$ 步里把两个数相加。假定可以在 $\log n$ 步里用 $n-1$ 个处理器的树形连接网络对 $n=2^k$ 个数求和。设 x_1，x_2，…，x_{2n}是希望求和的 $2n=2^{k+1}$个数。$2n-1$ 个处理器的树形连接网络包括 $n-1$ 个处理器的树形连接网络，以及对每个树叶作为儿子的两个新处理器。在一步之内，可以用这个较大的网络的树叶来求出 x_1+x_2，x_3+x_4，…，$x_{2n-1}+x_{2n}$，结果得出 n 个数，根据归纳假设，可以用网络的其余部分在 $\log n$ 步内求它们的和。因为使用了 $\log n+1$ 步而且 $\log(2n)=\log 2+\log n=1+\log n$，所以这样就完成了证明。

39. 只有 c　　41. c 和 h

43. 假设树 T 有至少两个中心。设 u 和 v 是不同的中心，都有离心度 e，u 和 e 不相邻。因为 T 是连通的，所以有从 u 到 v 的简单通路 P。设 c 是这个通路上的任意顶点。因为 c 的离心度至少为 e，所以存在顶点 w 使得从 c 到 w 的唯一简单通路长度至少为 e。显然这条铁路不能同时包含 u 和 v，否则将有一条简单回路。事实上，一旦这条从 c 到 w 的通路可能沿着 P 的一部分向 u 或 v 前进，它就离开 P 并且不返回 P。不失一般性，假定这条通路不沿着 P 向 u 前进。于是从 u 到 c 到 w 的通路是简单的，而且长度不超过 e，矛盾。因此 u 和 v 是相邻的。现在因为任何两个中心都是相邻的，所以假如有多于两个的中心，那么 T 就包含简单回路 K_3 作为子图，这是矛盾。

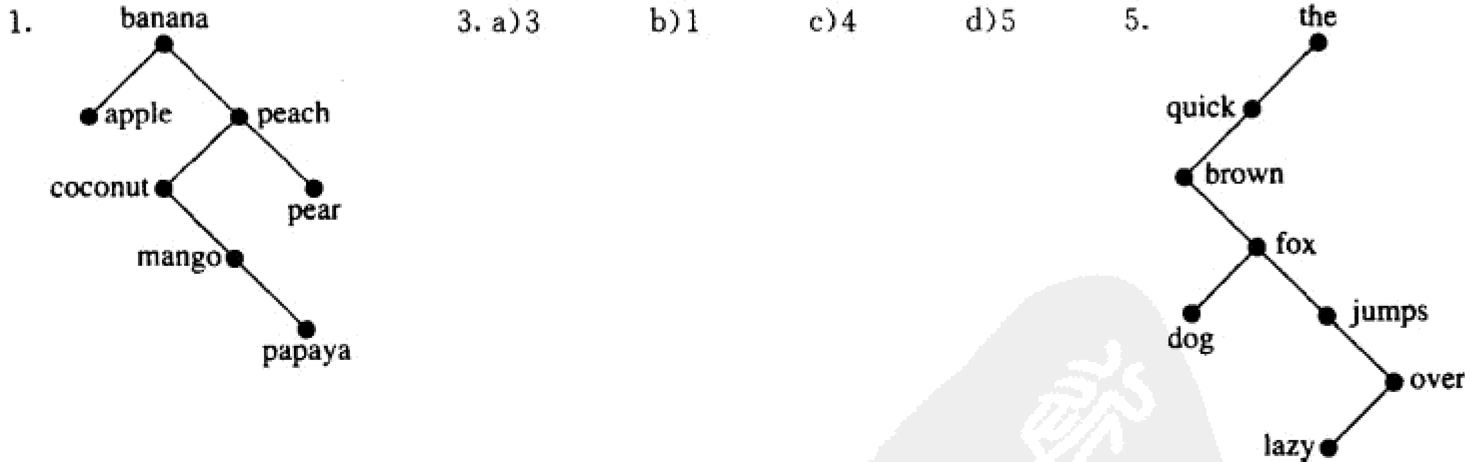
45.



47. 命题称有 n 个顶点的"每个"树都有长度为 $n-1$ 的路径。而证明只示出存在某个 n 个顶点的树有长度为 $n-1$ 的路径。

9.2 节

1.



3. a)3　　b)1　　c)4　　d)5　　5.

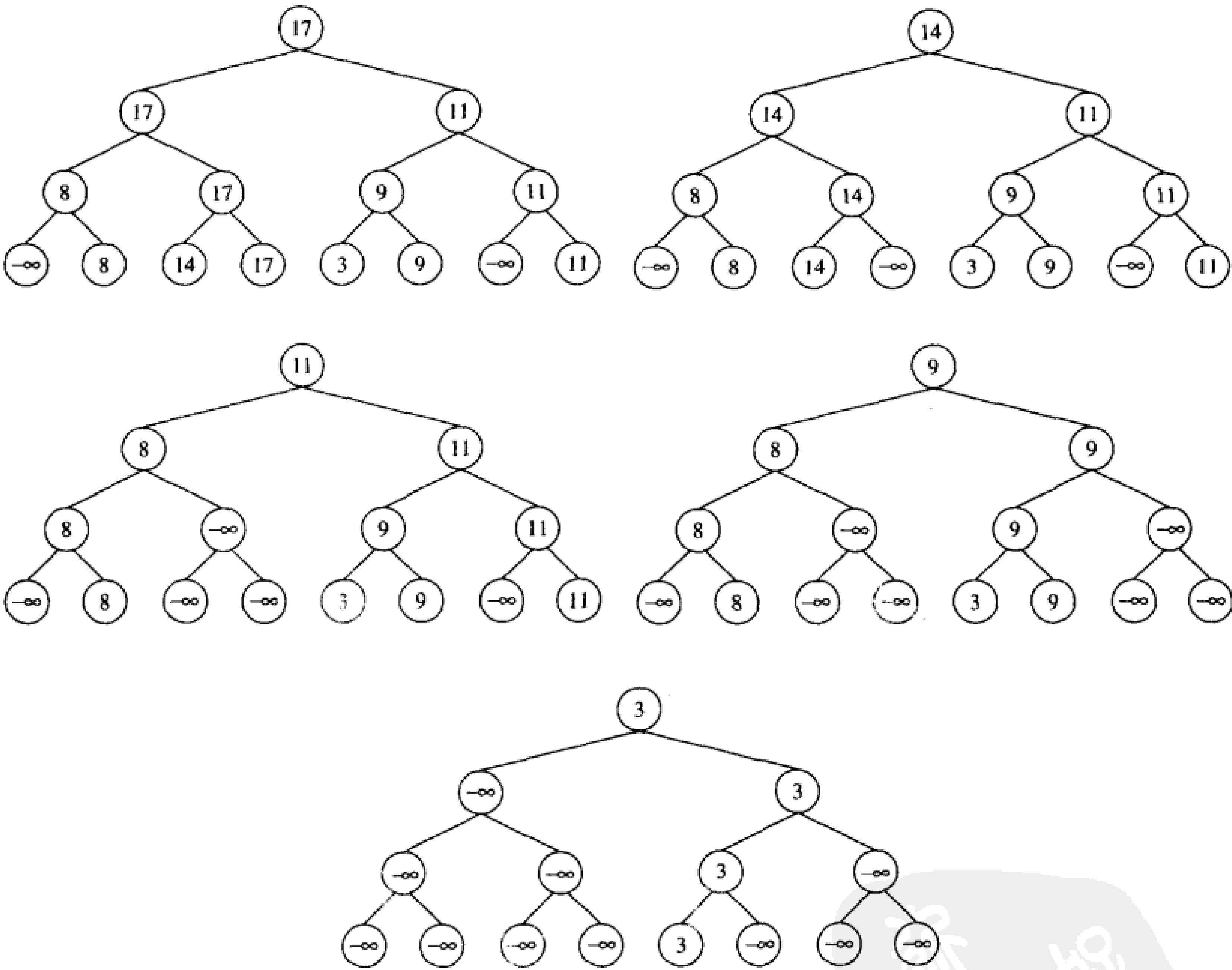
7. 至少需要 $\lceil \log_3 4 \rceil=2$ 次称重，因为只有四种结果(因为不要求确定硬币是较轻还是较重)。事实上，两次称重是足够的。首先称重硬币 1 和硬币 2。若它们平衡，则称重硬币 1 和硬币 3。若硬币 1 与硬币 3 重量相同，则硬币 4 是伪币，若它们重量不相同，则硬币 3 是伪币。若硬币 1 与硬币 2 重量不相同，则再称重硬币 1 和硬币 3。若它们平衡，则硬币 2 是伪币；若它们不平衡，则硬币 1 是伪币。

9. 至少需要 $\lceil \log_3 13 \rceil=3$ 次称重。事实上，三次称重是足够的。首先把硬币 1，2 和 3 放在天平的左边，而把

硬币 4，5 和 6 放在天平的右边。若相等，则应用例 2 到硬币 1，2，7，8，9，10，11 和 12 上。若不相等，则应用例 2 到硬币 1，2，3，4，5，6，7 和 8 上。

11. 最少比较 5 次。称这些元素为 a，b，c，d。先比较 a 和 b，再比较 c 和 d。不失一般性，假设 $a<b$ 和 $c<d$。再比较 a 和 c，其中较小的那个就是最小元素。仍不失一般性，假设 $a<c$。最后让 b 与 c 和 d 比较，就完全确定它们的顺序。

13. 前两步如课本所叙。在确定了 22 是第二大元素之后，在树中用 $-\infty$ 代替树叶 22，并沿着 22 上升到树根所经过的路径从树叶开始重新计算胜者。于是看到 17 是第三大元素，所以重复上述过程：在树中用 $-\infty$ 代替树叶 17 并重新计算。于是看到 14 是第四大元素，所以重复上述过程：在树中用 $-\infty$ 代替树叶 14 并重新计算。于是看到 11 是第五大元素，所以重复上述过程：在树中用 $-\infty$ 代替树叶 11 并重新计算。就这样继续这个过程。确定 9 是第六大元素，8 是第七大元素，3 是第八大元素。除了倒数第 2 步以外，所有其余步骤所产生的树如下图所示。



15. 顶点的值是目前在那里的表元素，顶点的标记是导致那个值的树叶的名称(即位置)。

procedure *tournament sort*$(a_1, \cdots, a_n)$

$k := \lceil \log n \rceil$

建立高度为 k 的二叉树

for $i := 1$ **to** n

　令第 i 个树叶的值是 a_i，令第 i 个树叶的标记为其自身

for $i := n+1$ **to** 2^k

　令第 i 个树叶的值是 $-\infty$，令第 i 个树叶的标记为其自身

for $i := k-1$ **downto** 0

```
      for 第 i 层上每个顶点 v
          令 v 的值是子女中的较大值，令 v 的标记为有较大值的那个子女的标记
  for i := 1 to n
  begin
      c_i := 根的值
      令 v 是根的标记
      令 v 的值是 −∞
      while 根的标记仍是 v
      begin
          v := parent(v)
          令 v 的值是子女中的较大值，令 v 的标记为有较大值的那个子女的标记
      end
  end{c_1, …, c_n 是非降序排列的表}
```

17. $k-1$，其中 $n=2^k$　　19. a)是　　b)否　　c)是　　d)是

21. a：000，e：001，i：01，k：1100，o：1101，p：11110，u：11111

23. a：11；b：101；c：100；d：01；e：00；2.25 位（注意：这个编码取决于如何打破平局，但平均位数总是一样的。）

25. 总共有四种可能的答案。这里给出一种，通过在这一种中交换 t 与 v 和（或）交换 u 与 w 来得到另外三种。

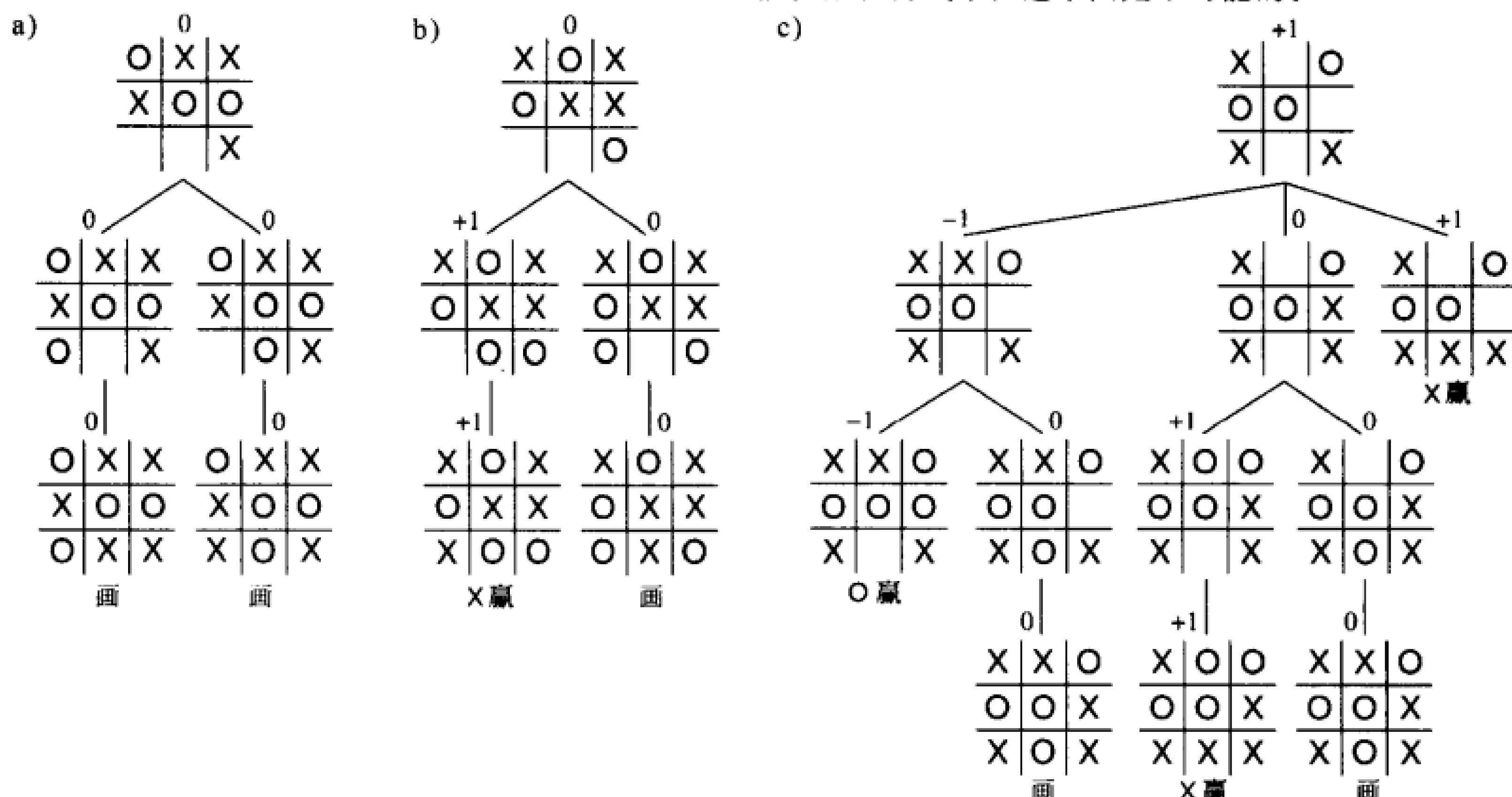
27. A：0001；B：101001；C：11001；D：00000；E：100；F：001100；G：001101；H：0101；I：0100；J：110100101；K：1101000；L：00001；M：10101；N：0110；O：0010；P：101000；Q：1101001000；R：1011；S：0111；T：111；U：00111；V：110101；W：11000；X：11010011；Y：11011；Z：1101001001。

29. A：2；E：1；N：010；R：011；T：02；Z：00。　　31. n

33. 由于这棵树比较大，所以在有些地方标明了"看课本"。参看图 9-24；以方框或圆圈顶点为根的子树与图 9-24 中的对应子树完全一样。先手获胜。

35. a) \$1　　b) \$3　　c) －\$3

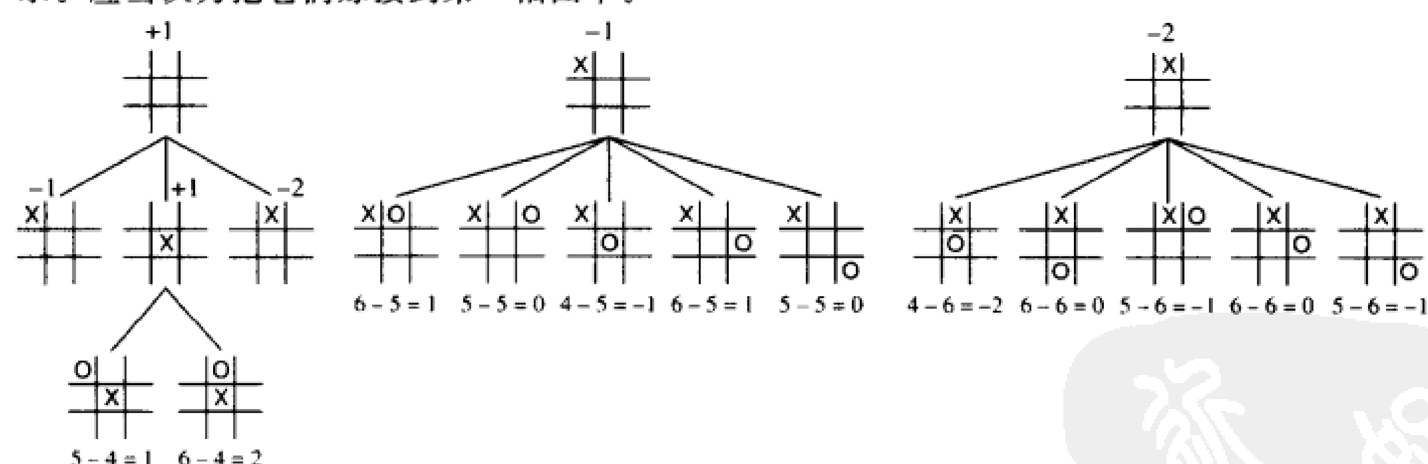
37. 参看下图。a)0　b)0　c)1　d)这种局面不可能发生在游戏中；这个图是不可能的。



39. 用强归纳法证明。基础步骤：当每堆有 $n=2$ 块石子时，若选手一从一堆取走两块石子，则选手二从剩下那堆取一块石子而获胜。若选手一从一堆取走一块石子，则选手二从另一堆取两块石子而获胜。归纳步骤：假设归纳假设，即对于所有 $2\leqslant j\leqslant k$，如果游戏从两堆各有 j 块的石子开始，则选手二获胜，其中 $k\geqslant 2$。考虑两堆各有 $k+1$ 块石子的游戏。若选手一从一堆取走所有石子，则选手二从剩下那堆取走除了一块外的所有石子而获胜。否则选手一在一堆中留下 j 块石子，其中 $2\leqslant j\leqslant k$，而在另一堆留下 $k+1$ 块石子。选手二从较大的一堆取走同样多石子，也给这堆留下 j 块石子。此时的游戏包含各有 j 块的两堆石子。根据归纳假设，在目前游戏中选手二(也是实际游戏中的选手二)获胜，强归纳法完成了。

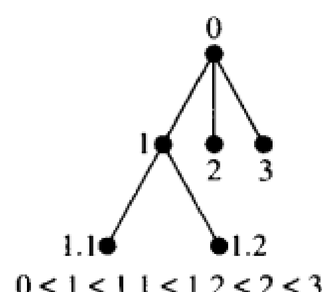
41. 7；49

43. 树的值是 1。注意：第二和第三棵树是第一棵树树根两个子女的子树，其子树由于篇幅限制而没有显示。应当认为把它们嫁接到第一幅图中。

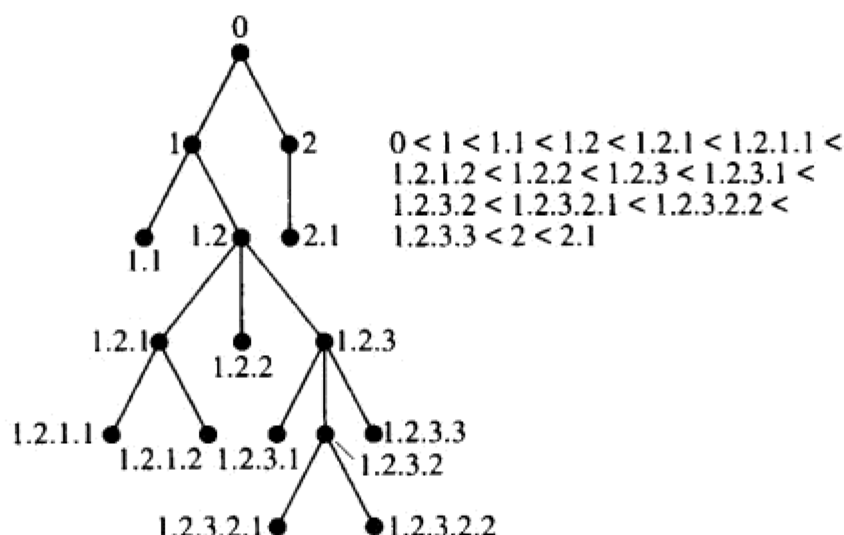


9.3 节

1.



3.



5. 否

7. a，b，d，e，f，g，c

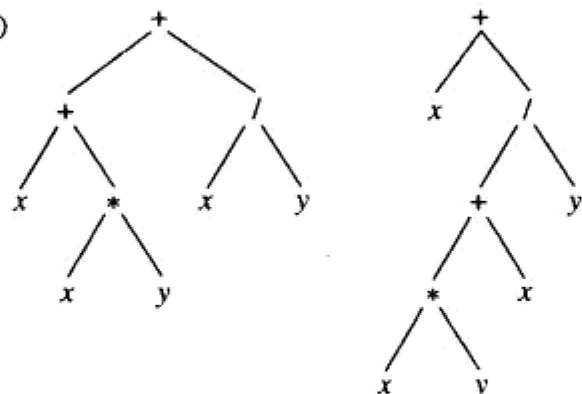
9. a，b，e，k，l，m，f，g，n，r，s，c，d，h，o，i，j，p，q

11. d，b，i，e，m，j，n，o，a，f，c，g，k，h，p，l

13. d，f，g，e，b，c，a

15. k，l，m，e，f，r，s，n，g，b，c，o，h，i，p，q，j，d，a

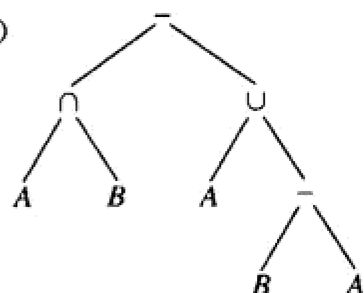
17. a)



b) $++x*xy/xy$，$+x/+*xyxy$　　c) $xxy*+xy/+$，$xxy*x+y/+$

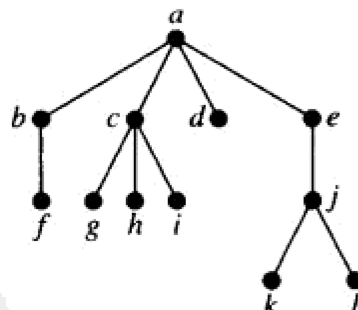
d) $((x+(x*y))+(x/y))$，$(x+(((x*y)+x)/y))$

19. a)



b) $-\cap AB\cup A-BA$　　c) $AB\cap ABA-\cup-$　　d) $((A\cap B)-(A\cup(B-A)))$

21. 14　　23. a) 1　　b) 1　　c) 4　　d) 2205　　25.



27. 用数学归纳法。对一个元素的表来说结果是平凡的。假定对 n 个元素的表来说结果为真。对于归纳步骤，从后面开始。找出表后面的顶点序列，从最后一个树叶开始，到根结束，每个顶点都是它后面那个顶点的最后一个儿子。删除这个树叶并且应用归纳假设。

29. 在每种情形里分别为 c，d，b，f，g，h，e，a。

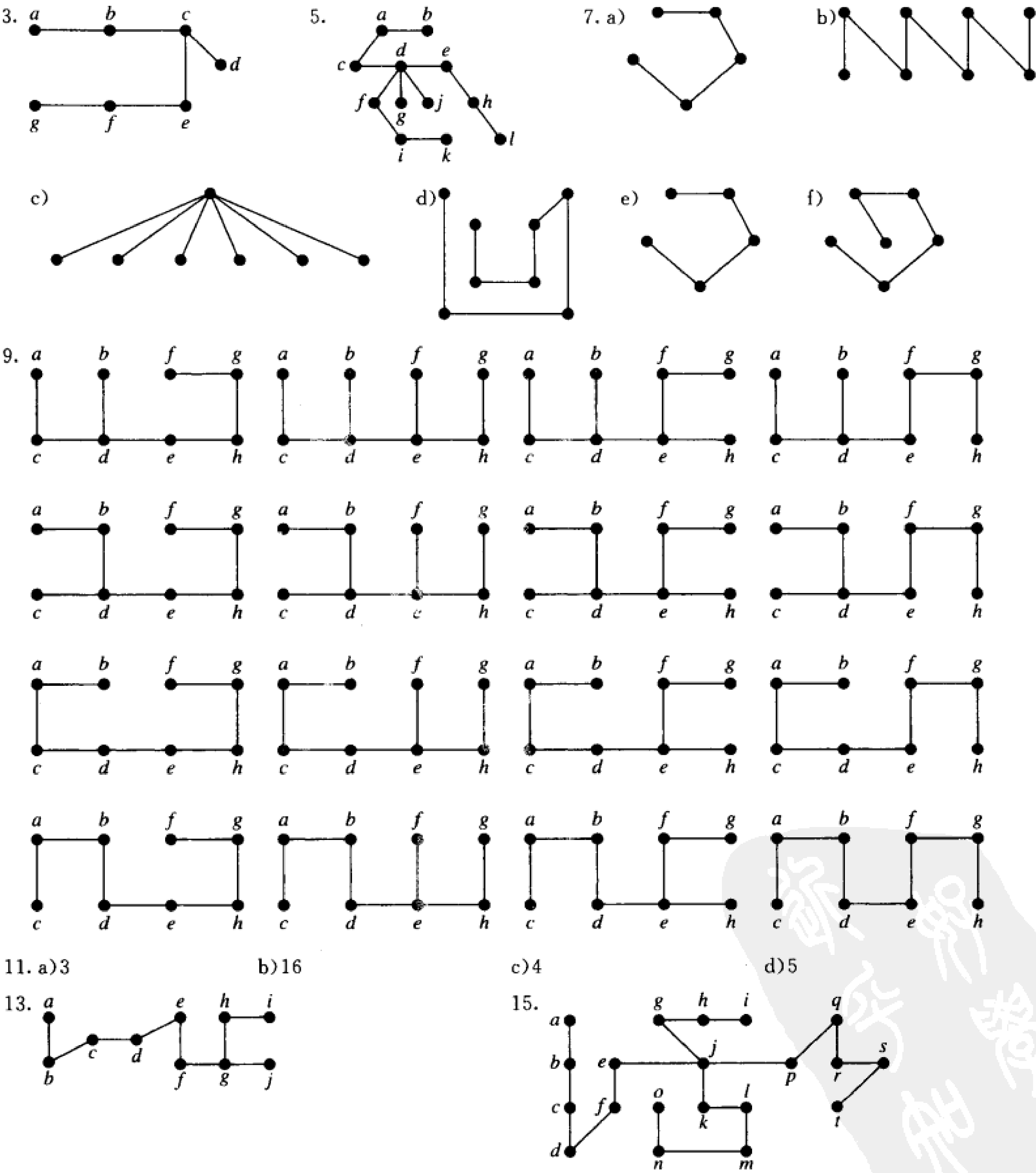
31. 用数学归纳法证明。设 $S(X)$ 和 $O(X)$ 分别表示合式公式 X 里的符号数和运算数。对长度为 1 的合式公

式来说命题为真，因为它们都有 1 个符号和 0 个运算。假定对长度小于 n 的合式公式来说命题为真。长度为 n 的合式公式必然形如 $*XY$，其中 $*$ 是运算而 X 和 Y 都是长度小于 n 的合式公式。于是根据归纳假设 $S(*XY)=S(X)+S(Y)=(O(X)+1)+(O(Y)+1)=O(X)+O(Y)+2$。因为 $O(*XY)=1+O(X)+O(Y)$，所以 $S(*XY)=O(*XY)+1$。

33. 例如 $xy+zx\circ+x\circ$，$xyz++yx++$，$xyxy\circ\circ xy\circ\circ z\circ+$，$xz\times zz+\circ$，$yyyy\circ\circ\circ$，$zx+yz+\circ$

9.4 节

1. $m-n+1$



11. a)3　　b)16　　c)4　　d)5

17. a)长度为 6 的路径　　b)长度为 5 的路径

c)长度为 6 的路径　　　d)与访问顶点的选择顺序有关；可能是长度为 7 的路径

19. 采用宽度优先搜索，初始顶点是中心顶点，处理该顶点时把 n 个轮辐顶点加入树中，因此产生的树是 $K_{1,n}$。采用深度优先搜索，从轮图的中心顶点开始，并访问一个相邻顶点(边缘上顶点中的一个)。从这个顶点转移到边缘上一个相邻顶点，继续这样绕圈下去，直到已经到达了每个顶点。因此产生的生成树是长度为 n 的路径。

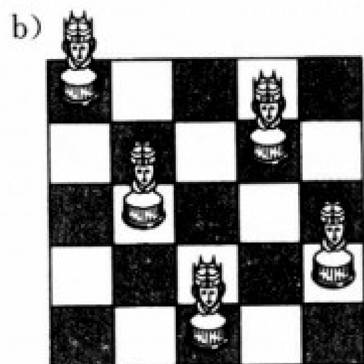
21. 采用宽度优先搜索，第一步从一个 m 度顶点辐射到所有 n 度顶点。下一步处理一个 n 度顶点，加入从这个顶点到所有剩余 m 度顶点的边。结果是一个 $K_{1,n-1}$ 和一个 $K_{1,m-1}$，中心用一边相连。采用深度优先搜索，从一边到另一边来回移动，直到不能前进为止。如果 $m=n$ 或 $m=n-1$，则得到长度为 $m+n-1$ 的路径。否则，该路径停在较大集合中的已经访问了一些顶点处，再返回到该路径上到顶点 v 的一个链接，然后，继续访问从顶点 v 出发的集合中的剩余顶点。

23. 一组可能不再继续下去的航班是：波士顿—纽约，底特律—波士顿，波士顿—华盛顿，纽约—华盛顿，纽约—芝加哥，亚特兰大—华盛顿，亚特兰大—达拉斯，亚特兰大—洛杉矶，亚特兰大—圣路易斯，圣路易斯—达拉斯，圣路易斯—底特律，圣路易斯—丹佛，达拉斯—圣迭哥，达拉斯—洛杉矶，达拉斯—旧金山，圣迭哥—洛杉矶，洛杉矶—旧金山，旧金山—西雅图。

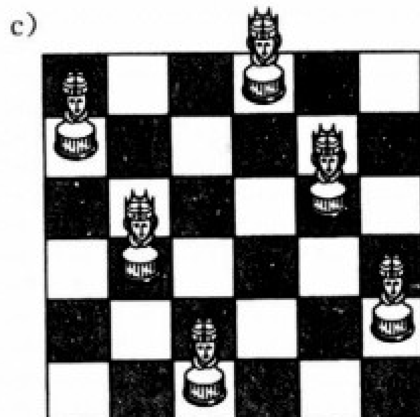
25. 树

27. 通过对通路的长度进行归纳来证明：若通路长度为 0，则结果是平凡的。若长度为 1，则 u 是与 v 相邻的，所以 u 是在宽度优先生成树的 1 层上。假定对长度为 l 的通路来说结果为真。若通路长度为 $l+1$，则设 u' 是从 v 到 u 的最短通路上倒数第二个顶点。根据归纳假设，u' 是在宽度优先生成树的 l 层上。假如 u 是在不超过 l 的层上，那么显然从 v 到 u 的最短通路的长度也不超过 l。所以在添加 l 层的顶点之后，仍然没有添加 u 到宽度优先生成树上。因为 u 是与 u' 相邻的，所以将把它添加在 $l+1$ 层上(尽管连接 u' 和 u 的边不必添加)。

29. a)无解　　b)



c)



31. 从某个顶点开始并且沿着一条通路前进，尽量不重复经过顶点，在所有顶点都已经访问过之后，允许返回到出发点。当不可能沿着一条通路继续下去时，就回溯并且尝试当前通路的另外一种扩充。

33. 取 G 的连通分支的生成树的并图。它们都是不相交的，所以结果是一个森林。

35. $m-n+c$

37. 在每个分支上使用深度优先搜索。

39. 如果在深度优先搜索过程中处理顶点 u 时，没有沿着边 uv 前进，则只能是这样的情形，就是已经访问过顶点 v 了。有两种情形。如果是在开始处理 u 之后访问的 v，则由于还没有结束处理 u，所以 v 一定出现在以 u 为根的子树中(因此一定是 u 的后代)。另一方面，如果是在开始处理 u 之前已经开始处理 v 了，那么为什么在那时没有沿着这条边前进呢？所以一定是还没有结束处理 v，换句话说，还在构造以 v 为根的子树，所以 u 是 v 的后代，因此 v 是 u 的祖先。

41. 如果正在处理的图本身是棵树，则这两个过程肯定产生相同的生成树，因为在这种情形下，只有一棵生成树(整个图)。而且这也是发生这种情况的唯一情形。如果原图有任何其他的边，则根据练习 39，这些边一定是后退边，因此把一个顶点连到祖先或后代，而根据练习 40，这些边一定连接同层上或相差一层上的顶点。显然这两种可能性是互斥的。因此如果两棵生成树注定是相同的，则除了树上的边

之外没有其他的边了。

43. 由于在这个过程中不沿着生成树以外的边前进，所以可以忽略这些边。因此可以假设这个图起初是一棵根树。基础情形是平凡的(只有一个顶点)，所以假设归纳假设，即应用在 n 个顶点的树上的宽度优先搜索按照顶点在树上的层次来访问这些顶点。考虑有 $n+1$ 个顶点的树 T。在这个树的宽度优先搜索中，最后一个访问的顶点是最后一个加入等待处理顶点表的那个顶点。当正在处理这个顶点的父母(比如说 u)时加入这个顶点。必须证明 v 是在树的最低(最底，即层数最大)的层上。假如不是这样，比如说顶点 x 是在更低层上，x 的父母是 w。则 w 是在比 u 更低的层上。显然 v 一定是树叶，因为在看见 v 之前不应该看见 v 的任何子女。考虑从 T 删除 v 后得到的树 T'。根据归纳假设，T' 中的顶点一定是按照其在 T' 中的层次来处理的(这个层次与其在 T 中的层次相同，T' 中少了 v 对算法其余部分没有影响)。因此一定在 w 之前处理过 u 了，因此 v 应当在 x 之前加入等待表，矛盾。因此 v 是在树的最底层，证毕。

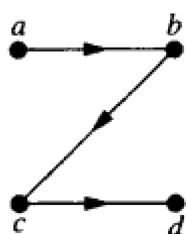
45. 这样修改算法 2 给出的伪代码：在算法开头把 m 初始化为 0，在从 L 删除 v 的语句后增加语句"$m := m+1$"和"把 m 赋给顶点 v"。

47. 如果在深度优先搜索过程中处理有向边 uv 的起点 u 时，没有沿着边 uv 前进，则只能是这样的情形，就是已经访问过顶点 v 了。有三种情形。如果是在开始处理 u 之后访问的 v，则由于还没有结束处理 u，所以 v 一定出现在以 u 为根的子树中(因此一定是 u 的后代)，所以有一条前进边。否则，在开始处理 u 之前一定已经开始处理 v 了。如果还没有结束处理 v(即还在构造以 v 为根的子树)，则 u 是 v 的后代，因此 v 是 u 的祖先(有一条后退边)。最后，如果已经结束了处理 v，则根据定义，就有了一条交叉边。

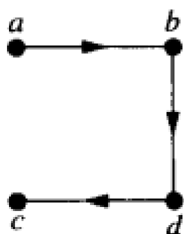
49. 设 T 是在图 9-41 里构造的生成树，而 T_1，T_2，T_3 和 T_4 是在图 9-42 里的生成树。用 $d(T', T'')$ 表示 T' 与 T'' 之间的距离。则 $d(T, T_1)=6$，$d(T, T_2)=4$，$d(T, T_3)=4$，$d(T, T_4)=2$，$d(T_1, T_2)=4$，$d(T_1, T_3)=4$，$d(T_1, T_4)=6$，$d(T_2, T_3)=4$，$d(T_2, T_4)=2$，$d(T_3, T_4)=4$。

51. 假定 $e_1=\{u, v\}$ 是像规定的那样。则 $T_2 \cup \{e_1\}$ 包含一个包含 e_1 的简单回路 C。图 $T_1-\{e_1\}$ 包含两个连通分支；e_1 的端点是在两个不同的分支里。从 u 开始按照与 e_1 相反的方向前进，直到你来到 v 所在的分支里第一个顶点为止。刚刚经过的边是 e_2。显然 $T_2 \cup \{e_1\}-\{e_2\}$ 是树，因为 e_2 是在 C 上。另外 $T_1-\{e_1\} \cup \{e_2\}$ 也是树，因为 e_2 重新连接这两个分支。

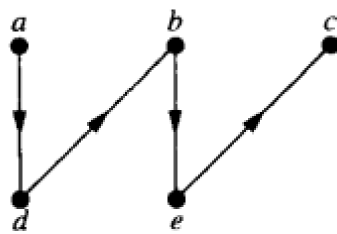
53. 练习 18：



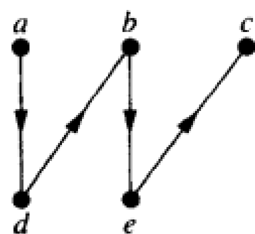
练习 19：



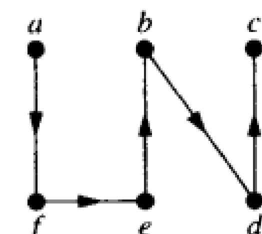
练习 20：



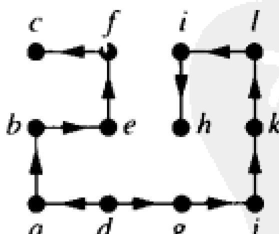
练习 21：



练习 22：



练习 23：



55. 首先构造这个有向图里的欧拉回路。然后从这个回路删除通向从前访问过的顶点的每条边。

57. 根据练习 56，一个有向图包含回路当且仅当存在后退边。可以如下检测出后退边。给每个顶点加标记指示其状态：未见到(初始情形)，已见到(进入 T)但未完成(即 visit(v)未终止)，完成(即 visit(v)已终止)。给算法 1 加入几行代码就能完成这个记录工作。于是为了确定一个有向图是否具有回路，只需要在查看边 uv 时检查 v 是否"已见到"。只要这样的事情一发生，就知道有回路；否则，就没有回路。

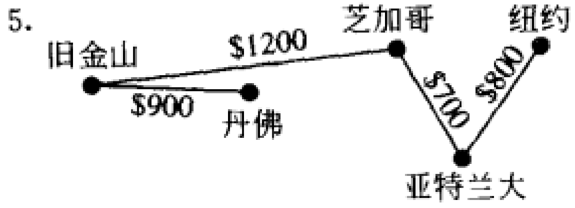
9.5 节

1. Deep Springs-Oasis，Oasis-Dyer，Oasis-Silverspeak，Silverspeak-Goldfield，Lida-Gold Point，Gold Point-

Beatty，Lida-Goldfield，Goldfield-Tonopah，Tonopah-Manhattan，Tonopah-Warm Springs

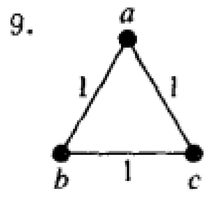
3. $\{e, f\}$，$\{c, f\}$，$\{e, h\}$，$\{h, i\}$，$\{b, c\}$，$\{b, d\}$，$\{a, d\}$，$\{g, h\}$

5.



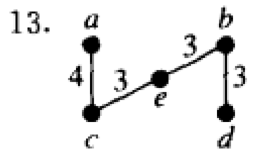
7. $\{e, f\}$，$\{a, d\}$，$\{h, i\}$，$\{b, d\}$，$\{c, f\}$，$\{e, h\}$，$\{b, c\}$，$\{g, h\}$

9.

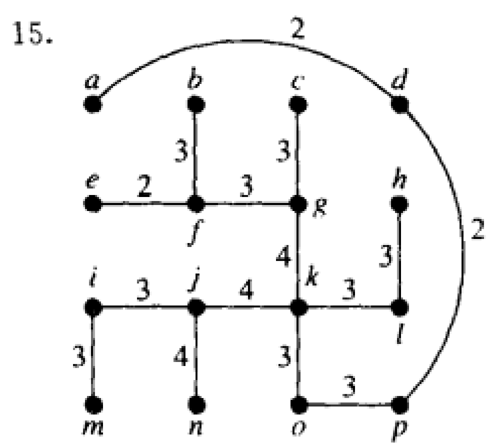


11. 在每阶段不选最小边，而是选性质相同的最大边。

13.



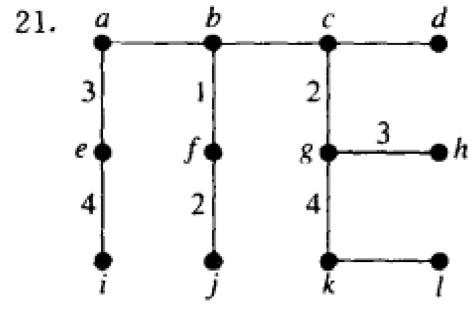
15.



17. 首先求 n 阶图 G 的最小生成树 T。然后对 $i=1$ 到 $n-1$，只从 G 中删除 T 的第 i 条边，并求剩余图的最小生成树。从这 $n-1$ 个树中挑选长度最短的一个。

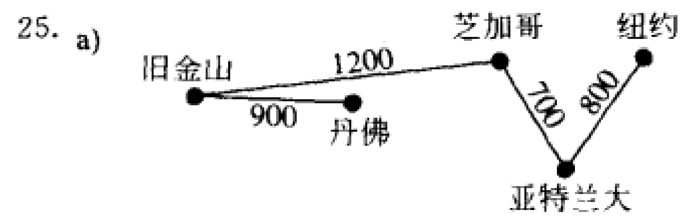
19. 如果所有边的大小都不同，在普林算法的正确性证明中，当把边 e_{k+1} 加入 T 并删除边 e 时就得出矛盾，而不是可能产生另一个生成树。

21.

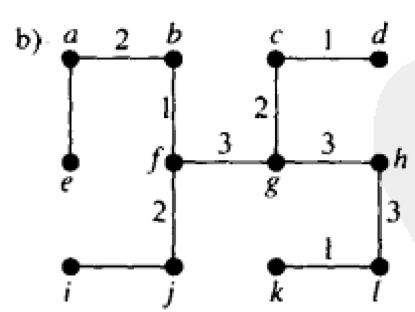


23. 与克鲁斯卡尔算法基本相同，不同的是，一开始 $T:=$ 指定的这组边，并且迭代是从 $i=1$ 到 $n-1-s$，s 是指定的这组边数。

25. a)



b)



27. 根据练习 24，索林算法每个阶段都产生森林。因此在选择 $n-1$ 条边后就产生树。还需证明这个树是最小生成树。设 T 是与索林树 S 具有最多公共边的最小生成树。如果 $T\neq S$，则存在边 $e\in S-T$，在算法的某个阶段 e 被加入，而在那个阶段之前，S 中所有的边也都在 T 中。$T\cup\{e\}$ 包含唯一的简单回路。在这个回路上找这样的边 $e'\in S-T$ 和 $e''\in T-S$，当把这个阶段的各个树看作"超顶点"时，e' 和 e'' 是"相邻的"。则根据这个算法，$w(e')\leqslant w(e'')$。于是把 T 换成 $T-\{e''\}\cup\{e'\}$ 就产生出比 T 更接近 S 的最小生成树。

29. 这 r 个树每一个都用一条新边至少连到一个其他的树上。因此结果中至多有 $r/2$ 个树(每个新树包含两个以上老树)。为了这样做，需要加入 $r-(r/2)=r/2$ 条边。因为加入边的数目是整数，所以至少是 $\lceil r/2\rceil$。

31. 如果 $k\geqslant \log n$，则 $n/2^k\leqslant 1$，故 $\lceil n/2^k\rceil=1$，所以根据练习 30，这个算法至多在 $\log n$ 次迭代后终止。

补充练习

1. 假设 T 是树，显然 T 没有简单回路。如果加入边 e 连接两个不相邻顶点 u 和 v，显然形成一个简单回路，因为把 e 加入 T 导致了图有太多的边而不再是树。形成的唯一简单回路包含边 e 和 T 中从 u 到 v 的唯一路径 P。假设 T 满足给定条件。只需证明 T 连通，因为图中没有简单回路。设 u 和 v 在不同连通分支里。加入边 $e=\{u, v\}$就不满足题目中的条件。

3. 假设树 T 有 n 个顶点，度分别为 $d_1, d_2, \cdots, d_n$，因为 $2e=\sum_{i=1}^{n} d_i$ 且 $e=n-1$，则有$2(n-1)=\sum_{i=1}^{n} d_i$。因为每个 $d_i\geqslant 1$，所以 $2(n-1)=n+\sum_{i=1}^{n}(d_i-1)$，即 $n-2=\sum_{i=1}^{n}(d_i-1)$。于是和中至多有 $n-2$ 项可以大于等于 1。故其中至少两项为 0。所以至少对两个 i 值有 $d_i=1$。

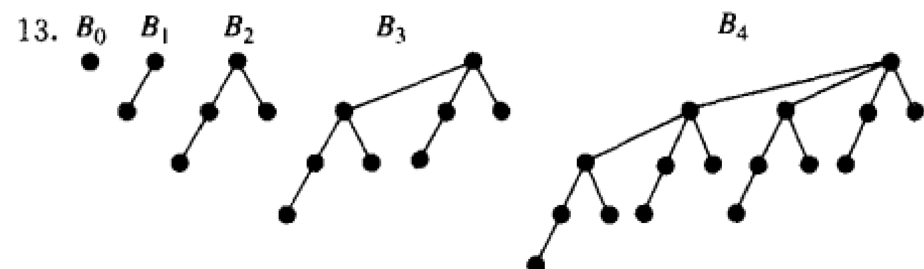
5. $2n-2$

7. T 没有回路，故不可能有子图同胚于 $K_{3,3}$ 或 K_5。

9. 分别给每个连通分支着色。对每个连通分支，首先确定树根，然后把所有偶数层顶点染成红色，把所有奇数层顶点染成蓝色。

11. 上界：k^k；下界：$2\lceil k/2\rceil^{k-1}$。

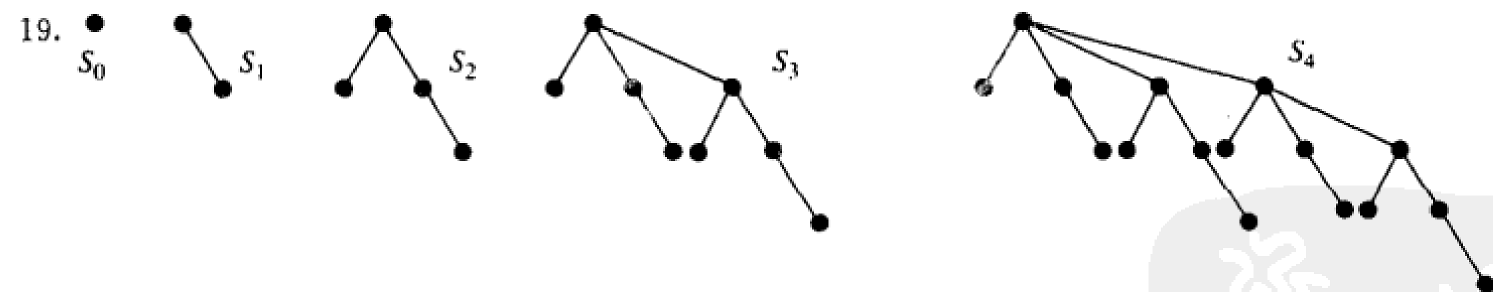
13.



15. 由于 B_{k+1} 由两个 B_k 构成，其中一个下移一层，故 k 每加 1 高度亦加 1。由于 B_0 高度为 0，经归纳得出 B_k 高度为 k。

17. 由于 B_{k+1} 的根是 B_k 的根外加一个子女(即另一个 B_k 的根)，故 k 每加 1 根的度亦加 1。由于 B_0 的根的度为 0，经归纳得出 B_k 的根的度为 k。

19.



21. 用数学归纳法。对 $k=0$ 结果为平凡的。假设结果对 $k-1$ 成立。T_{k-1} 是 T 的母树。根据归纳法，T 的子树可从 $T_0, \cdots, T_{k-2}$ 按所述方式得到。r_{k-2} 与 r_{k-1} 的最终相连是按照 S_k 树的定义中的描述。

23. **procedure** *level*(T：带根 r 的有序根树)

```
queue := 只含根 r 的序列
while queue 至少还含有一项
begin
    v := 队列中的第一个顶点
    列出 v
    从队列中删除 v 并把 v 的子女加入队尾
```

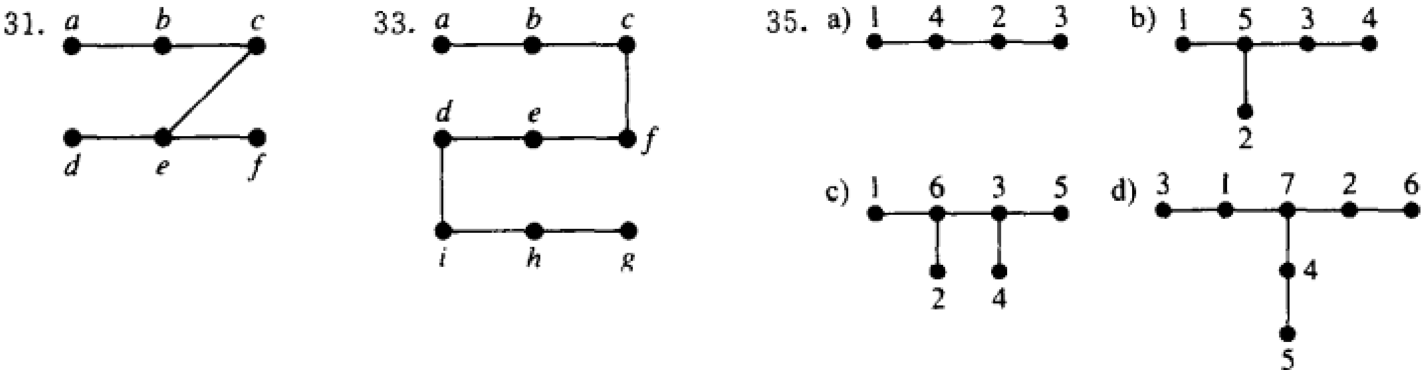
end

25. 这样构造树：对地址 0，插入树根。对每个带标记 i(i 是正整数)的顶点，插入子树，这个子树是从每个带标记 i，j(j 是正整数)的顶点的子树构造出来的。

27. a)是　　b)否　　c)是

29. 得到的图没有边是在超过一个的所述类型的简单回路上，所以是仙人掌图。

31. 33. 35.



37. 6

39. (a)0 表示 00，11 表示 01，100 表示 10，101 表示 11(具体编码与如何消除平局有关，但所有编码都是等价的)；对长度为 n 的串平均长度是 $0.645n$。(b)0 表示 000，100 表示 001，101 表示 010，110 表示 100，11100 表示 011，11101 表示 101，11110 表示 110，11111 表示 111(具体编码与如何消除平局有关，但所有编码都是等价的)；对长度为 n 的串平均长度是 $0.532\overline{6}n$。

41. 设从 G 删除顶点 v 和 v 关联的所有边后得到图 G'。可以这样得到 G 的最小生成树：取 v 关联的最小边和 G' 的一个最小生成树。

43. 假设边 e 是顶点 v 关联的最小边，假设 T 是不含 e 的最小生成树。把 e 加入 T，从因此形成的简单回路中删除与 v 关联的另一条边。将得到严格更小的一个生成树(因为删除的边比 e 大)。这是矛盾，所以 T 必含 e。

第 10 章

10.1 节

1. a)1　　b)1　　c)0　　d)0

3. a)

x	y	z	$\bar{x}y$
1	1	1	0
1	1	0	0
1	0	1	0
1	0	0	0
0	1	1	1
0	1	0	1
0	0	1	0
0	0	0	0

b)

x	y	z	$x+yz$
1	1	1	1
1	1	0	1
1	0	1	1
1	0	0	1
0	1	1	1
0	1	0	0
0	0	1	0
0	0	0	0

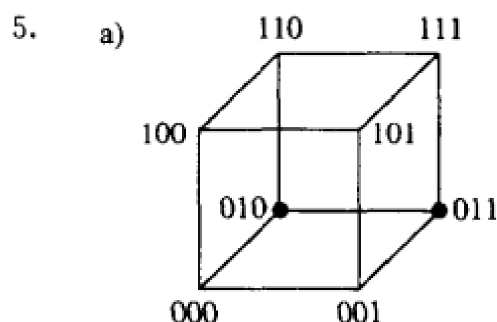
c)

x	y	z	$x\bar{y}+\overline{xyz}$
1	1	1	0
1	1	0	1
1	0	1	1
1	0	0	1
0	1	1	1
0	1	0	1
0	0	1	1
0	0	0	1

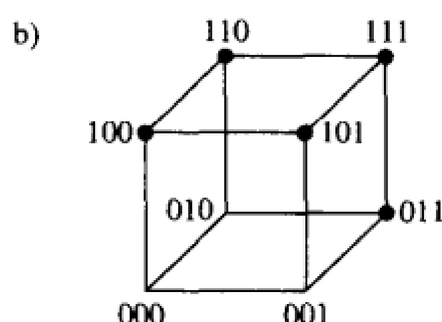
d)

x	y	z	$x(yz+\bar{y}\,\bar{z})$
1	1	1	1
1	1	0	0
1	0	1	0
1	0	0	1
0	1	1	0
0	1	0	0
0	0	1	0
0	0	0	0

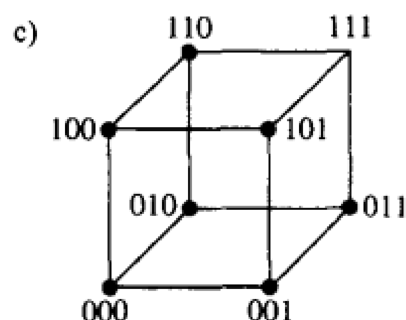
5. a)



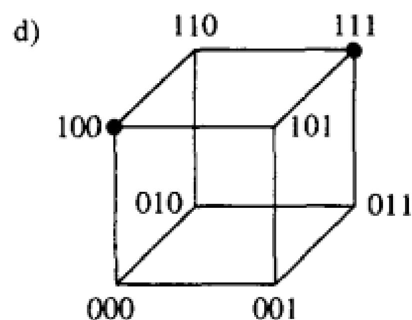
b)



c)



d)



7. (0，0)和(1，1)

9. $x+xy=x\cdot 1+xy=x(1+y)=x(y+1)=x\cdot 1=x$

11.

x	y	z	$x\bar{y}$	$y\bar{z}$	$\bar{x}z$	$x\bar{y}+y\bar{z}+\bar{x}z$	$\bar{x}y$	$\bar{y}z$	$x\bar{z}$	$\bar{x}y+\bar{y}z+x\bar{z}$
1	1	1	0	0	0	0	0	0	0	0
1	1	0	0	1	0	1	0	0	1	1
1	0	1	1	0	0	1	0	1	0	1
1	0	0	1	0	0	1	0	0	1	1
0	1	1	0	0	1	1	1	0	0	1
0	1	0	0	1	0	1	1	0	0	1
0	0	1	0	0	1	1	0	1	0	1
0	0	0	0	0	0	0	0	0	0	0

13.

x	$x+x$	$x\cdot x$
0	0	0
1	1	1

15.

x	$x+1$	$x\cdot 0$
0	1	0
1	1	0

17.

x	y	z	$y+z$	$x+(y+z)$	$x+y$	$(x+y)+z$	yz	$x(yz)$	xy	$(xy)z$
1	1	1	1	1	1	1	1	1	1	1
1	1	0	1	1	1	1	0	0	1	0
1	0	1	1	1	1	1	0	0	0	0
1	0	0	0	1	1	1	0	0	0	0
0	1	1	1	1	1	1	1	0	0	0
0	1	0	1	1	1	1	0	0	0	0
0	0	1	1	1	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0

19.

x	y	xy	$(\overline{xy})$	$\bar{x}$	$\bar{y}$	$\bar{x}+\bar{y}$	$x+y$	$(\overline{x+y})$	$\bar{x}\bar{y}$
1	1	1	0	0	0	0	1	0	0
1	0	0	1	0	1	1	1	0	0
0	1	0	1	1	0	1	1	0	0
0	0	0	1	1	1	1	0	1	1

21. $0\cdot\bar{0}=0\cdot 1=0$；$1\cdot\bar{1}=1\cdot 0=0$

23.

x	y	$x\oplus y$	$x+y$	xy	$(\overline{xy})$	$(x+y)(\overline{xy})$	$x\bar{y}$	$\bar{x}y$	$x\bar{y}+\bar{x}y$
1	1	0	1	1	0	0	0	0	0
1	0	1	1	0	1	1	1	0	1
0	1	1	1	0	1	1	0	1	1
0	0	0	0	0	1	0	0	0	0

25. a)对，用真值表可证。　b)错，如取 $x=1$，$y=1$，$z=1$。　c)错，如取 $x=1$，$y=1$，$z=0$。

27. 根据德摩根律，对一个表达式的求补，除了变元取补外，其余如同求此表达式的对偶。

29. 16

31. 由控制律、分配律和同一律，$x\vee x=(x\vee x)\wedge 1=(x\vee x)\wedge(x\vee\bar{x})=x\vee(x\wedge\bar{x})=x\vee 0=x$。类似地，$x\wedge x=(x\wedge x)\vee 0=(x\wedge x)\vee(x\wedge\bar{x})=x\wedge(x\vee\bar{x})=x\wedge 1=x$。

33. 由同一律和交换律，$0\vee 1=1$ 且 $0\wedge 1=0$，从而 $\bar{0}=1$。类似地，由于 $1\vee 0=1$ 且 $1\wedge 0=0$，从而 $\bar{1}=0$。

35. 首先注意 $x\wedge 0=0$ 且 $x\vee 1=1$ 对任意 x 成立，这很容易证明。为证明第一个等式，只要证明$(x\vee y)\vee(\bar{x}\wedge\bar{y})=1$ 且$(x\vee y)\wedge(\bar{x}\wedge\bar{y})=1$。由结合律、交换律、分配律、控制律和同一律知：$(x\vee y)\vee(\bar{x}\wedge\bar{y})=y\vee(x\vee(\bar{x}\wedge\bar{y}))=y\vee((x\vee\bar{x})\wedge(x\vee\bar{y}))=y\vee(1\wedge(x\vee\bar{y}))=y\vee(x\vee\bar{y})=(y\vee\bar{y})\vee x=1\vee x=1$，且$(x\vee y)\wedge(\bar{x}\wedge\bar{y})=\bar{y}\wedge(\bar{x}\wedge(x\vee y))=\bar{y}\wedge((\bar{x}\wedge x)\vee(\bar{x}\wedge y))=\bar{y}\wedge(0\vee(\bar{x}\wedge y))=\bar{y}\wedge(\bar{x}\wedge y)=\bar{x}\wedge(y\wedge\bar{y})=\bar{x}\wedge 0=0$。类似可证第二个等式。

37. 由假设、练习 31 和分配律可得：$x=x\vee 0=x\vee(x\vee y)=(x\vee x)\vee y=x\vee y=0$。类似有 $y=0$。为证明第二个等式，注意 $x=x\wedge 1=x\wedge(x\wedge y)=(x\wedge x)\wedge y=x\wedge y=1$。类似地，$y=1$。

39. 用第 7 章补充练习 39 和 41 以及有补分配格的定义来建立定义中的五对定律。

10.2 节

1. a)$\bar{x}\bar{y}z$　b)$\bar{x}y\bar{z}$　c)$\bar{x}yz$　d)$\bar{x}\bar{y}\bar{z}$

3. a)$xyz+xy\bar{z}+x\bar{y}z+x\bar{y}\bar{z}+\bar{x}yz+\bar{x}y\bar{z}+\bar{x}\bar{y}z$　b)$xyz+xy\bar{z}+\bar{x}yz$
c)$xyz+xy\bar{z}+x\bar{y}z+x\bar{y}\bar{z}$　d)$x\bar{y}z+x\bar{y}\bar{z}$

5. $wxy\bar{z}+wx\bar{y}z+w\bar{x}yz+\bar{w}xyz+\bar{w}x\bar{y}\bar{z}+\bar{w}\bar{x}\bar{y}z+\bar{w}\bar{x}y\bar{z}+w\bar{x}\bar{y}\bar{z}$

7. a)$\bar{x}+\bar{y}+z$　b)$x+y+z$　c)$x+\bar{y}+z$

9. $y_1+y_2+\cdots y_n=0$ 当且仅当 $y_i=0$ 对 $i=1, 2, \cdots, n$ 都成立。本题成立当且仅当：若 $y_i=x_i$，则 $x_i=0$；若 $y_i=\bar{x}_i$，则 $x_i=1$。

11. a)$x+y+z$
b)$(x+y+z)(x+y+\bar{z})(x+\bar{y}+z)(\bar{x}+y+z)(\bar{x}+y+\bar{z})$
c)$(x+y+z)(x+y+\bar{z})(x+\bar{y}+z)(x+\bar{y}+\bar{z})$
d)$(x+y+z)(x+y+\bar{z})(x+\bar{y}+z)(x+\bar{y}+\bar{z})(\bar{x}+\bar{y}+z)(\bar{x}+\bar{y}+\bar{z})$

13. a)$x+y+z$
b)$x+\overline{(y+\overline{(\bar{x}+z)})}$
c)$\overline{(x+\bar{y})}$
d)$\overline{(x+\overline{(x+\bar{y}+\bar{z})})}$



15. a)

x	$\overline{x}$	$x \downarrow x$
1	0	0
0	1	1

b)

x	y	xy	$x \downarrow x$	$y \downarrow y$	$(x \downarrow x) \downarrow (y \downarrow y)$
1	1	1	0	0	1
1	0	0	0	1	0
0	1	0	1	0	0
0	0	0	1	1	0

c)

x	y	$x+y$	$(x \downarrow y)$	$(x \downarrow y) \downarrow (x \downarrow y)$
1	1	1	0	1
1	0	1	0	1
0	1	1	0	1
0	0	0	1	0

17. a) $(((x \mid x) \mid (y \mid y)) \mid ((x \mid x) \mid (y \mid y))) \mid (z \mid z)$

b) $(((x \mid x) \mid (z \mid z)) \mid y) \mid (((x \mid x) \mid (z \mid z)) \mid y)$

c) x

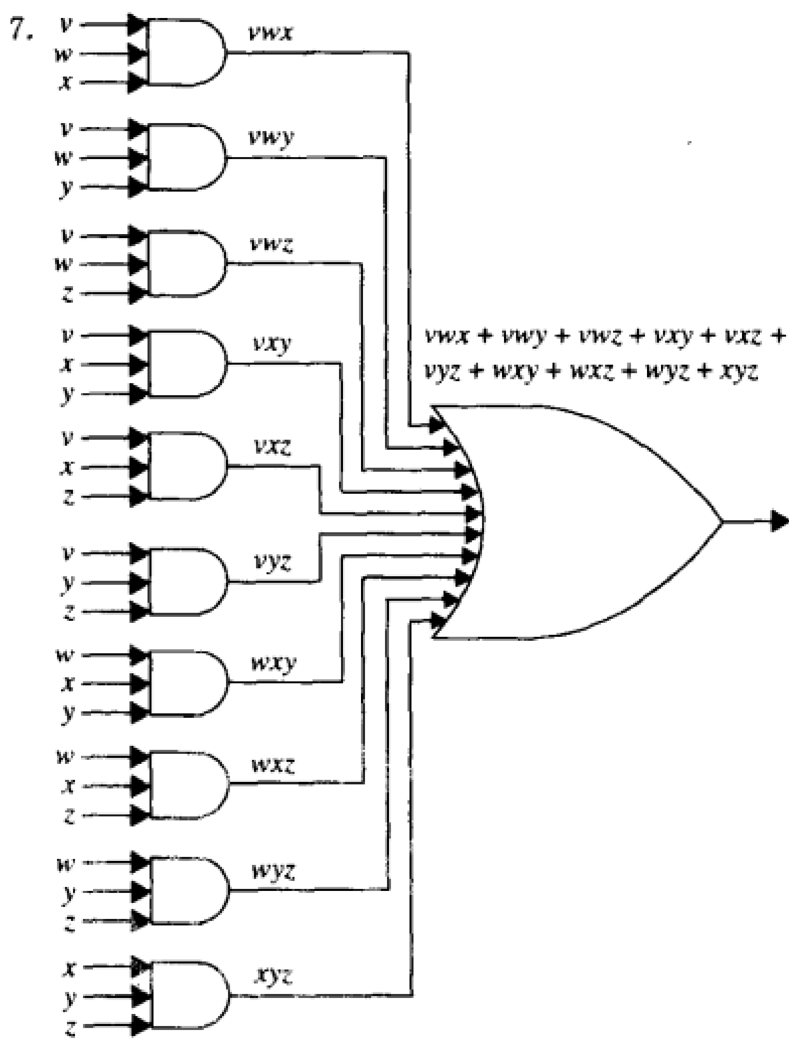
d) $(x \mid (y \mid y)) \mid (x \mid (y \mid y))$

19. 用+和·不能表示$\overline{x}$，因为当输入为1时，无法取得值0。

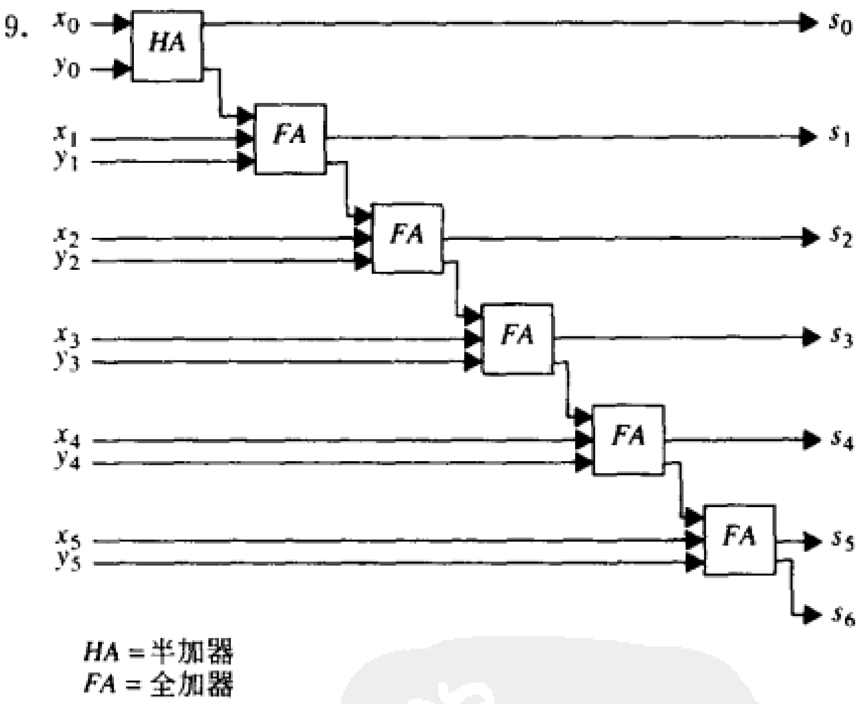
10.3 节

1. $(x+y)\overline{y}$ 3. $(\overline{xy})+(\overline{z}+x)$ 5. $(x+y+z)+(\overline{x}+y+z)+(\overline{x}+\overline{y}+\overline{z})$

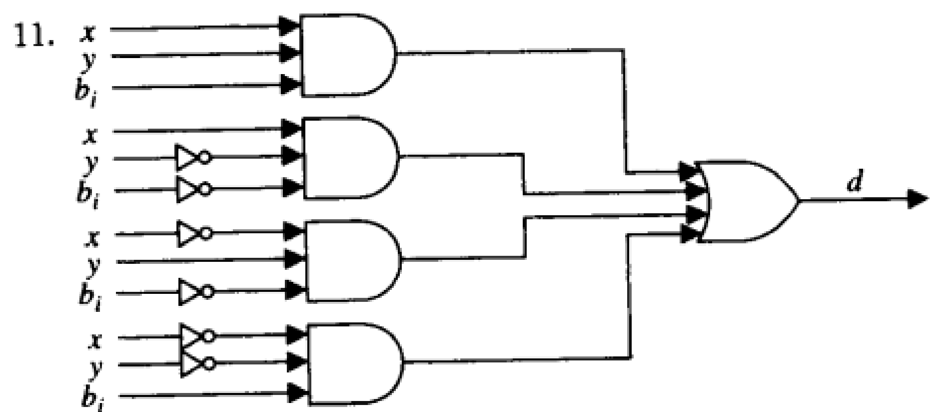
7.

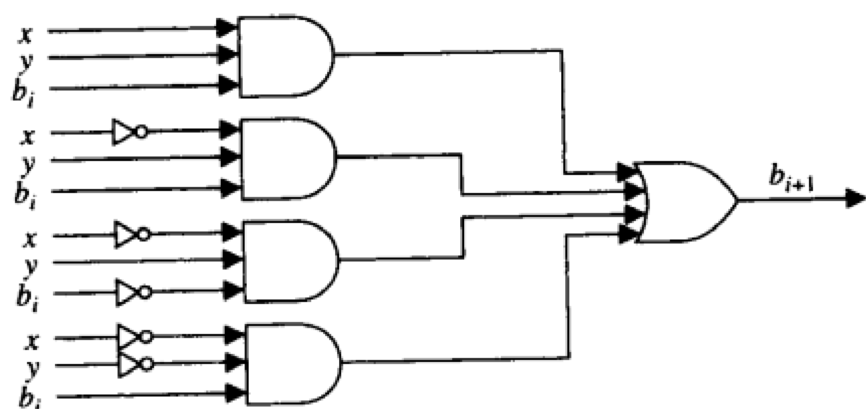


9.

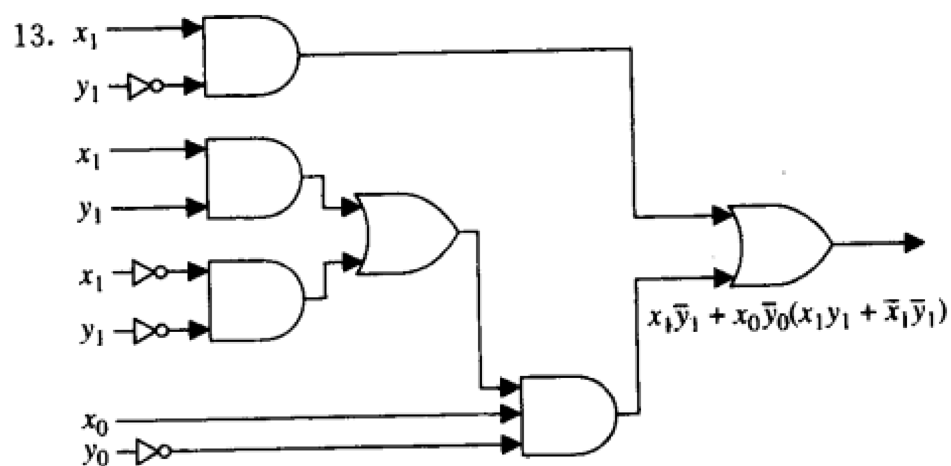


11.

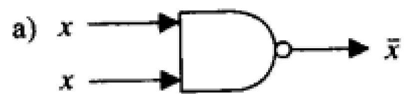

x
y
b_i
x
y
b_i
x
y
b_i
x
y
b_i
d

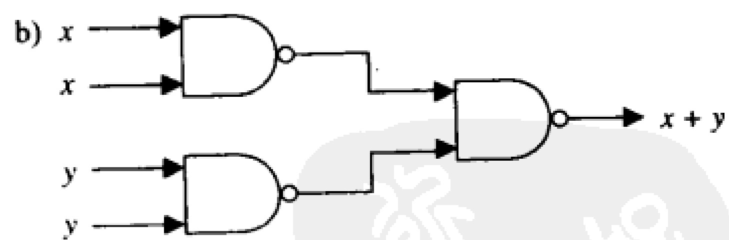

x
y
b_i
x
y
b_i
x
y
b_i
x
y
b_i
b_{i+1}

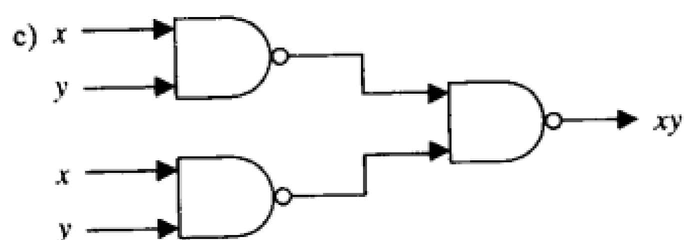
13.

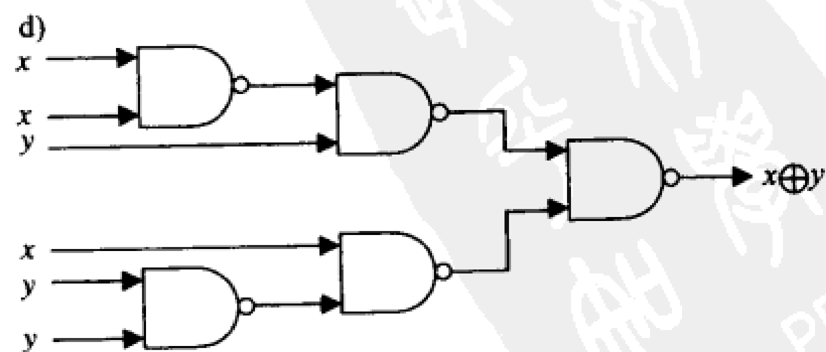

x_1
y_1
x_1
y_1
x_1
y_1
x_0
y_0
$x_1\overline{y}_1 + x_0\overline{y}_0(x_1y_1 + \overline{x}_1\overline{y}_1)$

15.

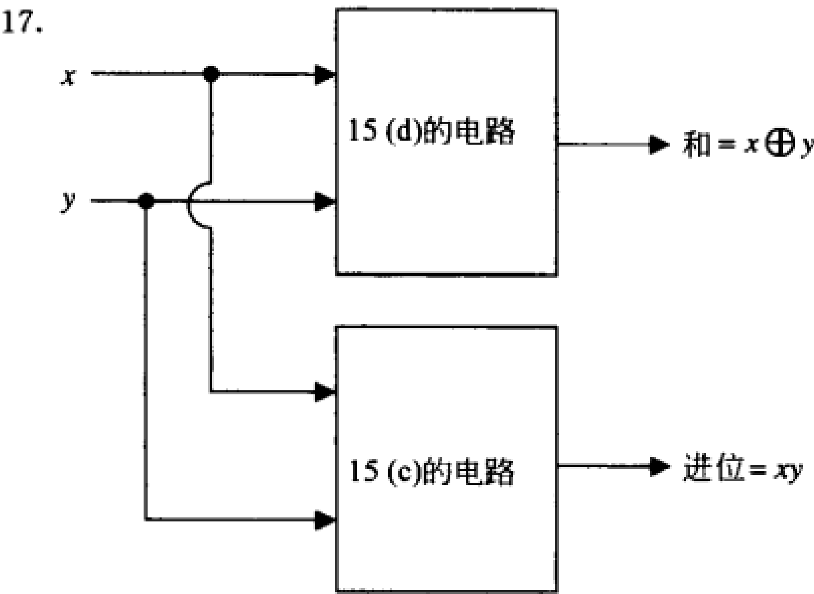

a) x
x
$\bar{x}$


b) x
x
y
y
x + y

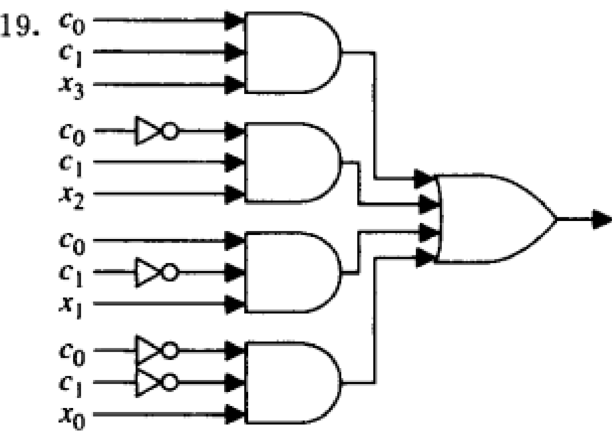

c) x
y
x
y
xy


d)
x
x
y
x
y
y
$x\oplus y$

17.



19.



10.4 节

1. a)

	y	$\overline{y}$
x		
$\overline{x}$	1	

b) xy 和 $\overline{xy}$

3. a)

	y	$\overline{y}$
x		1
$\overline{x}$		

b)

	y	$\overline{y}$
x	1	
$\overline{x}$		1

c)

	y	$\overline{y}$
x	1	1
$\overline{x}$	1	1

5. a)

	yz	$y\overline{z}$	$\overline{y}\,\overline{z}$	$\overline{y}z$
x				
$\overline{x}$		1		

b) $\overline{x}yz$，$\overline{x}\,\overline{y}\,\overline{z}$，$xy\overline{z}$

7. a)

	yz	$y\overline{z}$	$\overline{y}\,\overline{z}$	$\overline{y}z$
x			1	
$\overline{x}$				

b)

	yz	$y\overline{z}$	$\overline{y}\,\overline{z}$	$\overline{y}z$
x				
$\overline{x}$	1		1	

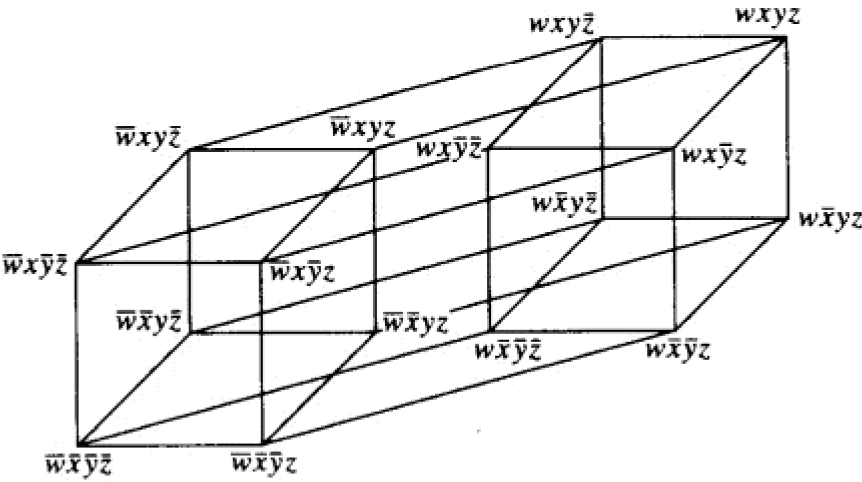
c)

	yz	$y\overline{z}$	$\overline{y}\,\overline{z}$	$\overline{y}z$
x	1	1		1
$\overline{x}$				1

9. 隐含数：xyz，$xy\overline{z}$，$x\overline{y}\,\overline{z}$，$\overline{x}y\overline{z}$，$xy$，$x\overline{z}$，$y\overline{z}$；素隐含数：$xy$，$x\overline{z}$，$y\overline{z}$；本原素隐含数：$xy$，$x\overline{z}$，$y\overline{z}$。

	yz	$y\overline{z}$	$\overline{y}\,\overline{z}$	$\overline{y}z$
x	1	1	1	
$\overline{x}$		1		

11. 右边的 3 立方体对应 w；整个图的顶表面给出的 3 立方体表示 x；背面给出的 3 立方体表示 y；左边和右边两个 3 立方体的右面给出的 3 立方体表示 z。在每种情形下，相对的三面表示互补的文字。表示 wz 的 2 立方体是右边的 3 立方体的右面；表示 $\overline{x}y$ 的 2 立方体是后边的 3 立方体的底面；表示 $\overline{y}\,\overline{z}$ 的 2 立方体是左前面。



13.

a)

	yz	$y\bar{z}$	$\bar{y}\bar{z}$	$\bar{y}z$
wx				
$w\bar{x}$				
$\bar{w}\bar{x}$				
$\bar{w}x$		1		

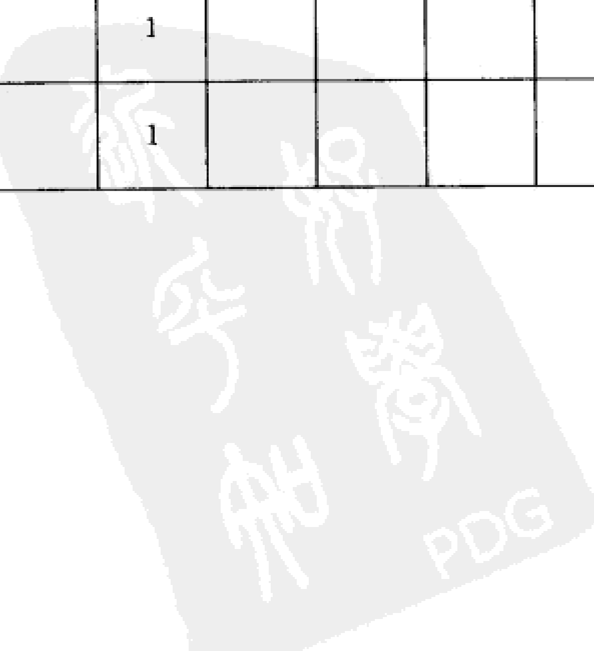
b) $\bar{w}xyz$, $\bar{w}\bar{x}y\bar{z}$, $\bar{w}x\bar{y}\bar{z}$, $\bar{w}x\bar{y}\bar{z}$

15. a)

	$x_3x_4x_5$	$x_3x_4\bar{x}_5$	$x_3\bar{x}_4\bar{x}_5$	$x_3\bar{x}_4x_5$	$\bar{x}_3\bar{x}_4x_5$	$\bar{x}_3\bar{x}_4\bar{x}_5$	$\bar{x}_3x_4\bar{x}_5$	$\bar{x}_3x_4x_5$
x_1x_2	1	1						
$x_1\bar{x}_2$								
$\bar{x}_1\bar{x}_2$								
$\bar{x}_1x_2$								

b)

	$x_3x_4x_5$	$x_3x_4\bar{x}_5$	$x_3\bar{x}_4\bar{x}_5$	$x_3\bar{x}_4x_5$	$\bar{x}_3\bar{x}_4x_5$	$\bar{x}_3\bar{x}_4\bar{x}_5$	$\bar{x}_3x_4\bar{x}_5$	$\bar{x}_3x_4x_5$
x_1x_2								
$x_1\bar{x}_2$								
$\bar{x}_1\bar{x}_2$	1			1				
$\bar{x}_1x_2$	1			1				



c)

	$x_3x_4x_5$	$x_3x_4\bar{x}_5$	$x_3\bar{x}_4\bar{x}_5$	$x_3\bar{x}_4x_5$	$\bar{x}_3\bar{x}_4x_5$	$\bar{x}_3\bar{x}_4\bar{x}_5$	$\bar{x}_3x_4\bar{x}_5$	$\bar{x}_3x_4x_5$
x_1x_2	1	1					1	1
$x_1\bar{x}_2$								
$\bar{x}_1\bar{x}_2$								
$\bar{x}_1x_2$	1	1					1	1

d)

	$x_3x_4x_5$	$x_3x_4\bar{x}_5$	$x_3\bar{x}_4\bar{x}_5$	$x_3\bar{x}_4x_5$	$\bar{x}_3\bar{x}_4x_5$	$\bar{x}_3\bar{x}_4\bar{x}_5$	$\bar{x}_3x_4\bar{x}_5$	$\bar{x}_3x_4x_5$
x_1x_2					1	1		
$x_1\bar{x}_2$					1	1		
$\bar{x}_1\bar{x}_2$					1	1		
$\bar{x}_1x_2$					1	1		

e)

	$x_3x_4x_5$	$x_3x_4\bar{x}_5$	$x_3\bar{x}_4\bar{x}_5$	$x_3\bar{x}_4x_5$	$\bar{x}_3\bar{x}_4x_5$	$\bar{x}_3\bar{x}_4\bar{x}_5$	$\bar{x}_3x_4\bar{x}_5$	$\bar{x}_3x_4x_5$
x_1x_2	1	1	1	1				
$x_1\bar{x}_2$	1	1	1	1				
$\bar{x}_1\bar{x}_2$	1	1	1	1				
$\bar{x}_1x_2$	1	1	1	1				

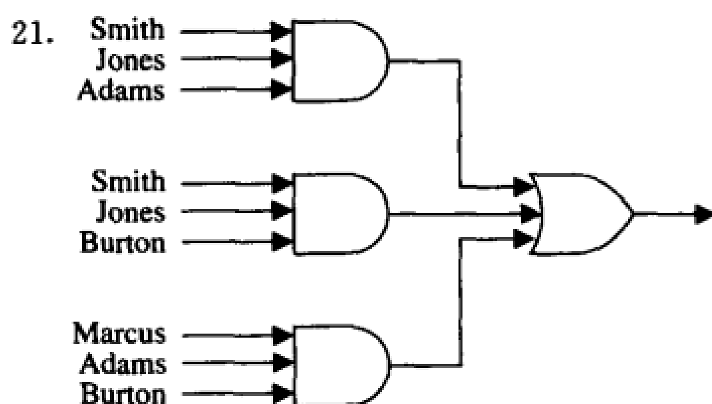
f)

	$x_3x_4x_5$	$x_3x_4\bar{x}_5$	$x_3\bar{x}_4\bar{x}_5$	$x_3\bar{x}_4x_5$	$\bar{x}_3\bar{x}_4x_5$	$\bar{x}_3\bar{x}_4\bar{x}_5$	$\bar{x}_3x_4\bar{x}_5$	$\bar{x}_3x_4x_5$
x_1x_2		1	1			1	1	
$x_1\bar{x}_2$		1	1			1	1	
$\bar{x}_1\bar{x}_2$		1	1			1	1	
$\bar{x}_1x_2$		1	1			1	1	

17. a)64 b)6

19. 行 1 和行 4 应该是相邻的。要考虑相邻性的列的对是：列 1 和 4，1 和 12，1 和 16，2 和 11，2 和 15，3 和 6，3 和 10，列 4 和 9，5 和 8，5 和 16，6 和 15，7 和 10，7 和 14，8 和 13，9 和 12，11 和 14，13 和 16。

21.



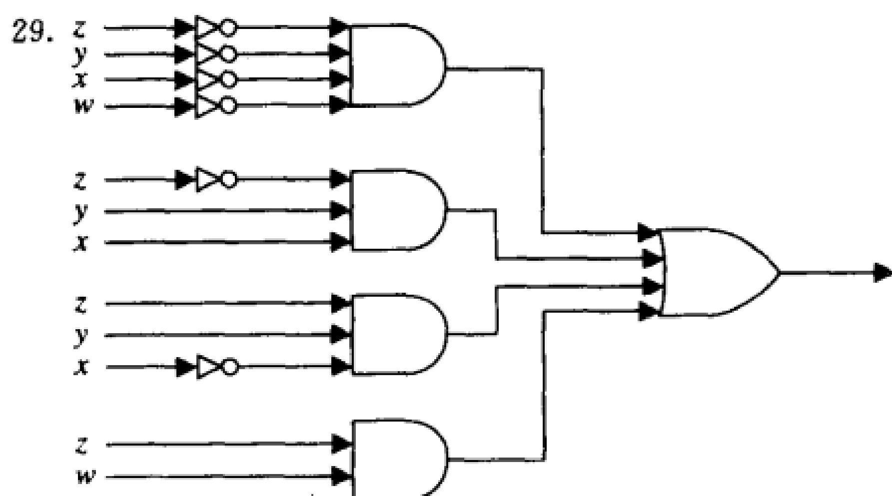
23. a) $\bar{x}z$　　b) y　　c) $x\bar{z}+\bar{x}z+\bar{y}z$　　d) $xz+\bar{x}y+\bar{y}\bar{z}$

25. a) $wxz+wx\bar{y}+w\bar{y}z+w\bar{x}y\bar{z}$　　b) $x\bar{y}z+\bar{w}\bar{y}z+wxy\bar{z}+w\bar{x}yz+\bar{w}\bar{x}y\bar{z}$

　c) $\bar{y}z+wxz+w\bar{x}\bar{y}+\bar{w}\bar{x}y\bar{z}$　　d) $wy+yz+\bar{x}y+wxz+\bar{w}\bar{x}z$

27. $x(y+z)$

29.



31. $\bar{x}\bar{z}+xz$

33. 对 n 应用归纳法。若 $n=1$，则考察一端标记 0 和另一端标记 1 的线段。k 只可能取的值也是 1，且若文字是 x_1，则我们所得的子立方体是由标记 1 的端点组成的 0 维子立方体，同时若文字是 $\bar{x}_1$，则所得的子立方体是由标记 0 的端点组成的 0 维子立方体。现假设命题对 n 为真；我们必须证明它对于 $n+1$ 为真。若文字 x_{n+1}（或者它的补）不是积的组成部分，则由归纳假设，当积在 n 个变量的取值中被观察到时，它对应于 n 维立方体的一个 $(n-k)$ 维子立方体，并且这个子立方体同线段[0, 1]的笛卡儿积给出维数高于我们所给 $(n+1)$ 维立方体的一个子立方体，即如所要求的那样具有 $(n+1)-k$ 维。另一方面，若文字 x_{n+1}（或者它的补）是积的组成部分，则剩余 $k-1$ 个文字的积对应于 n 维立方体中的一个 $n-(k-1)=(n+1)-k$ 维子立方体，并且在最后变量中的 1 端或者 0 端的片段是要求的子立方体。

补充练习

1. a) $x=0$，$y=0$，$z=0$；$x=1$，$y=1$，$z=1$。

　b) $x=0$，$y=0$，$z=0$；$x=0$，$y=0$，$z=1$；$x=0$，$y=1$，$z=0$；$x=1$，$y=0$，$z=1$；$x=1$，$y=1$，$z=0$；$x=1$，$y=1$，$z=1$。

　c) 无值。

3. a) 是　　b) 不是　　c) 不是　　d) 是　　5. 2^{n-1}

7. a) 若 $F(x_1, \cdots, x_n)=1$，由控制律得：$(F+G)(x_1, \cdots, x_n)=F(x_1, \cdots, x_n)+G(x_1, \cdots, x_n)=1$。故 $F\leqslant F+G$。

　b) 若 $(FG)(x_1, \cdots, x_n)=1$，则 $F(x_1, \cdots, x_n)\cdot G(x_1, \cdots, x_n)=1$，故 $F(x_1, \cdots, x_n)=1$。由此得 $FG\leqslant F$。

9. 因为 $F(x_1, \cdots, x_n)=1$ 蕴含 $F(x_1, \cdots, x_n)=1$，故 $\leqslant$ 是自反的。设 $F\leqslant G$ 且 $G\leqslant F$，则 $F(x_1, \cdots, x_n)=1$

当且仅当 $G(x_1, \cdots, x_n)=1$，这蕴含 $F=G$，故≤是反对称的。设 $F \leqslant G \leqslant H$，则由 $F(x_1, \cdots, x_n)=1$ 可得 $G(x_1, \cdots, x_n)=1$，此蕴含 $H(x_1, \cdots, x_n)=1$，从而 $F \leqslant H$，故≤是传递的。

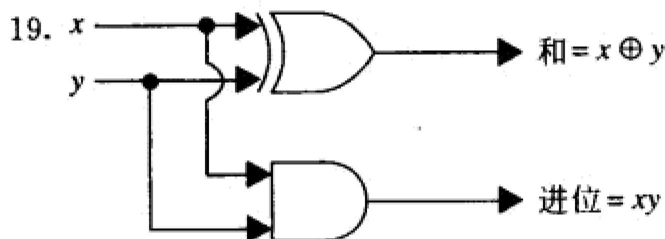
11. a) $x=1$，$y=0$，$z=0$ b) $x=1$，$y=0$，$z=0$ c) $x=1$，$y=0$，$z=0$

13.

x	y	$x \odot y$	$x \oplus y$	$\overline{(x \oplus y)}$
1	1	1	0	1
1	0	0	1	0
0	1	0	1	0
0	0	1	0	1

15. 是，如真值表所示。

17. a) 6 b) 5 c) 5 d) 6

19. 

21. $x_3 + x_2 \bar{x}_1$

23. 设它具有权 a 和 b，则存在实数 T 使得对 (1, 0) 和 (0, 1)，有 $xa+yb \geqslant T$；但对 (0, 0) 和 (1, 1)，有 $xa+yb<T$。从而 $a \geqslant T$，$b \geqslant T$，$0<T$，且 $a+b<T$。这样 a 和 b 都是正的，这蕴含 $a+b>a \geqslant T$，矛盾。

第 11 章

11.1 节

1. a) 句子⇒名词短语 不及物动词短语⇒冠词 形容词 名词 不及物动词短语⇒冠词 形容词 名词 不及物动词⇒…(3 步后)…⇒*the happy hare runs*。

   b) 句子⇒名词短语 不及物动词短语⇒冠词 形容词 名词 不及物动词短语⇒冠词 形容词 名词 不及物动词 副词…(4 步后)…⇒*the sleepy tortoise runs quickly*。

   c) 句子⇒名词短语 及物动词短语 名词短语⇒冠词 名词 及物动词短语 名词短语⇒冠词 名词 及物动词 名词短语⇒冠词 名词 及物动词 冠词 名词⇒…(4 步后)…⇒*the tortoise passes the hare*。

   d) 句子⇒名词短语 及物动词短语 名词短语⇒冠词 形容词 名词 及物动词短语 名词短语⇒冠词 形容词 名词 及物动词 名词短语⇒冠词 形容词 名词 及物动词 冠词 形容词 名词⇒…(6 步后)…⇒*the sleepy hare passes the happy tortoise*。

3. 使得末尾是名词(如 tortoise)只有一个方法，就是将一个名词短语放在末尾，这只能用如下产生式实现：句子→名词短语 及物动词短语 名词短语。而及物动词短语→及物动词→*passes*，但这个句子不含 *passes*。

5. $S \Rightarrow 0S1 \Rightarrow 00S11 \Rightarrow 000S111 \Rightarrow 000111$。

7. a) $S \Rightarrow 0S \Rightarrow 00S \Rightarrow 00S1 \Rightarrow 00S11 \Rightarrow 00S111 \Rightarrow 00S1111 \Rightarrow 001111$。

   b) $S \Rightarrow 0S \Rightarrow 00S \Rightarrow 001A \Rightarrow 0011A \Rightarrow 00111A \Rightarrow 001111$。

9. $S \Rightarrow 0SAB \Rightarrow 00SABAB \Rightarrow 00ABAB \Rightarrow 00AABB \Rightarrow 001ABB \Rightarrow 0011BB \Rightarrow 00112B \Rightarrow 001122$。

11. a) $S \rightarrow 00S$，$S \rightarrow \lambda$。

    b) $S \rightarrow 10A$，$A \rightarrow 00A$，$A \rightarrow \lambda$。

    c) $S \rightarrow AAS$，$S \rightarrow BBS$，$AB \rightarrow BA$，$BA \rightarrow AB$，$S \rightarrow \lambda$，$A \rightarrow 0$，$B \rightarrow 1$。

    d) $S \rightarrow 0000000000A$，$A \rightarrow 0A$，$A \rightarrow \lambda$。

e)$S\to AS$，$S\to ABS$，$S\to A$，$AB\to BA$，$BA\to AB$，$A\to 0$，$B\to 1$。

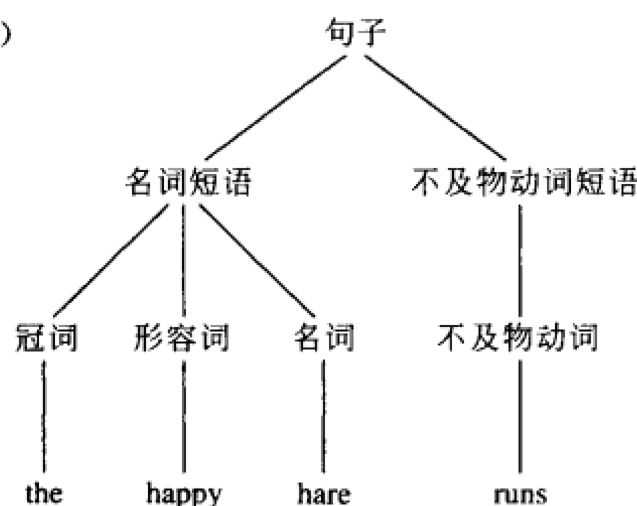
f)$S\to ABS$，$S\to \lambda$，$AB\to BA$，$BA\to AB$，$A\to 0$，$B\to 1$。

g)$S\to ABS$，$S\to T$，$S\to U$，$T\to AT$，$T\to A$，$U\to BU$，$U\to B$，$AB\to BA$，$BA\to AB$，$A\to 0$，$B\to 1$。

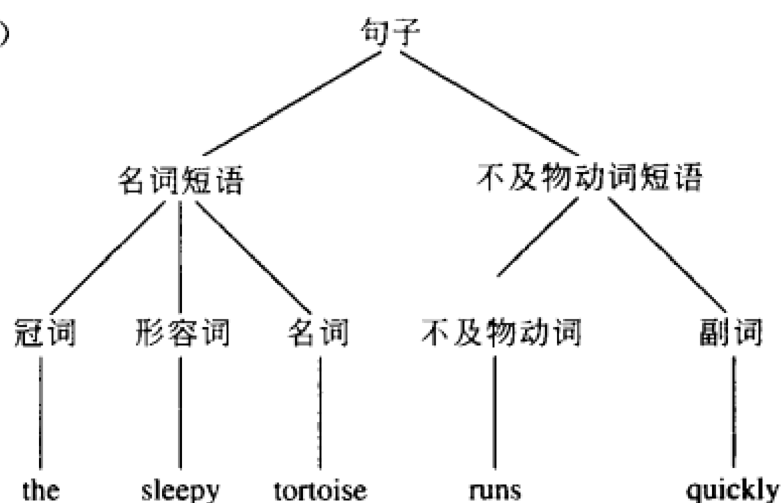
13. a)2 型，非 3 型　　b)3 型，非 2 型　　c)0 型，非 1 型
d)2 型，非 3 型　　e)2 型　　f)0 型，非 1 型
g)3 型　　h)0 型，非 1 型　　i)2 型，非 3 型
j)2 型，非 3 型

15. 设 S_1 和 S_2 分别为 G_1 和 G_2 的初始符，S 是一个新初始符。
a)增加 S 和产生式 $S\to S_1$ 及 $S\to S_2$。
b)增加 S 和产生式 $S\to S_1S_2$。
c)增加 S 和产生式 $S\to\lambda$ 及 $S\to S_1S$。

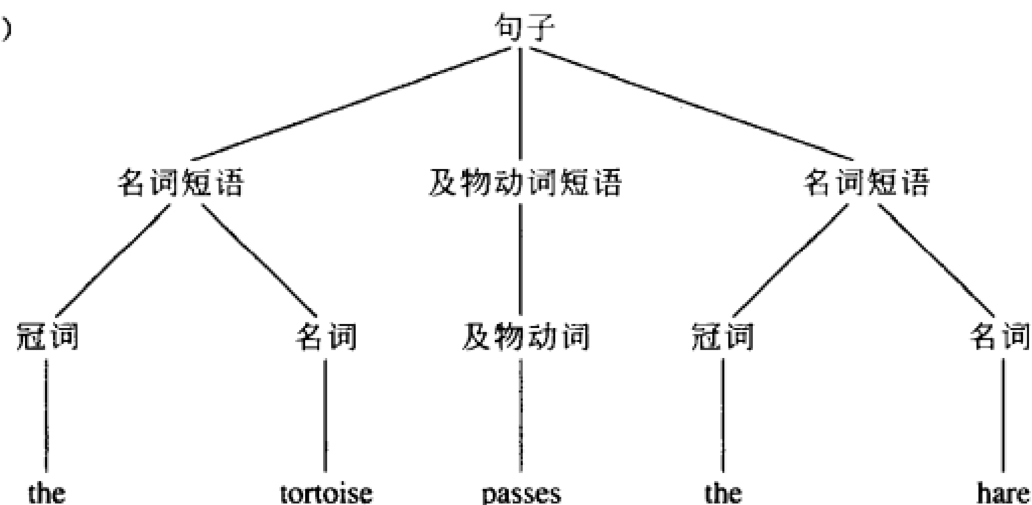
17. a)



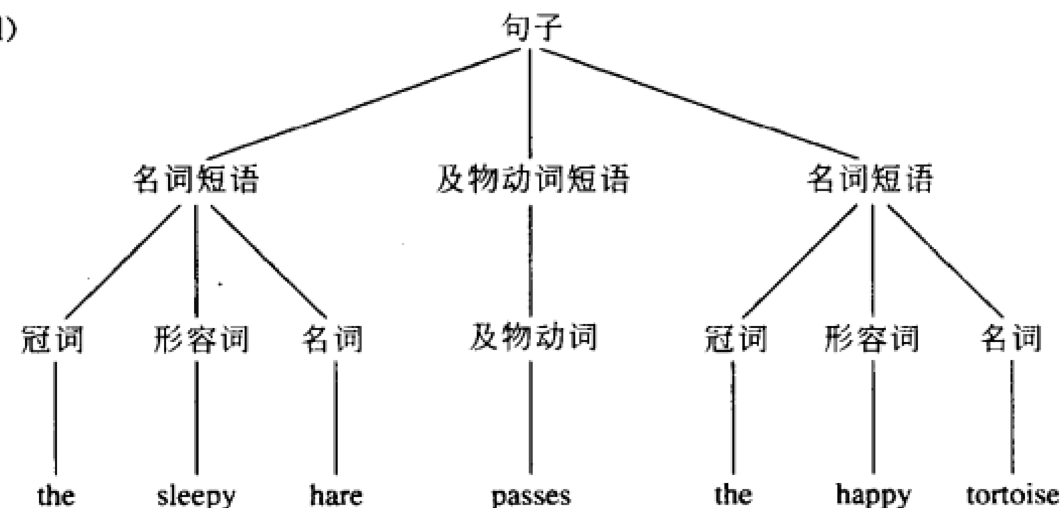
b)



c)

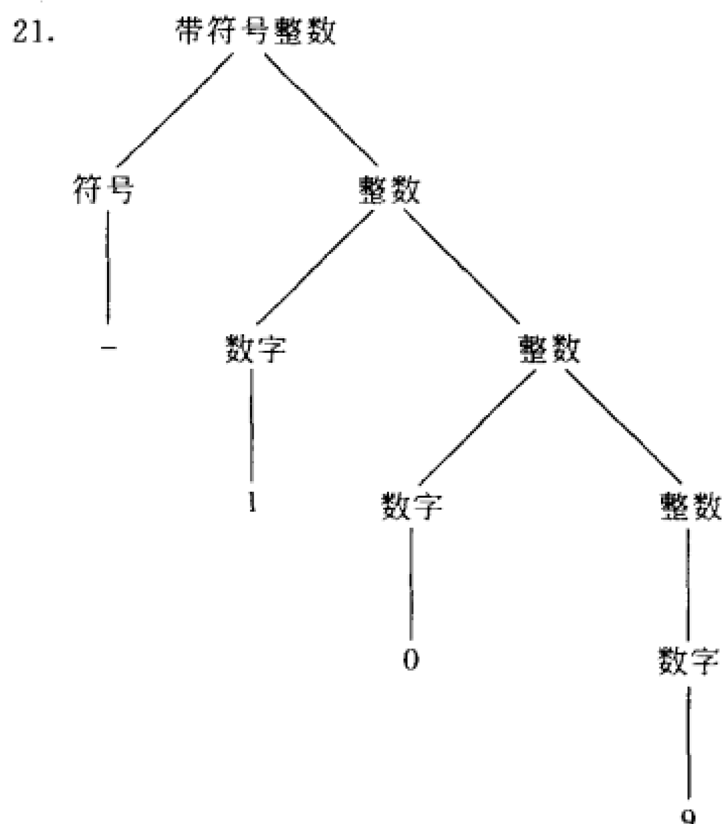


d)



19. a)是　　b)不是　　c)是　　d)不是

21.



23. a) S→〈符号〉〈整数〉

S→〈符号〉〈整数〉.〈正整数〉

〈符号〉→+

〈符号〉→−

〈整数〉→〈数字〉

〈整数〉→〈整数〉〈数字〉

〈数字〉→i，i=1，2，3，4，5，6，7，8，9，0

〈正整数〉→〈整数〉〈非 0 数字〉

〈正整数〉→〈非 0 数字〉〈整数〉

〈正整数〉→〈整数〉〈非 0 数字〉〈整数〉

〈正整数〉→〈非 0 数字〉

〈非 0 数字〉→i，i=1，2，3，4，5，6，7，8，9

b)〈带符号十进制数〉∷=〈符号〉〈整数〉|〈符号〉〈整数〉.〈正整数〉

〈符号〉∷=+|−

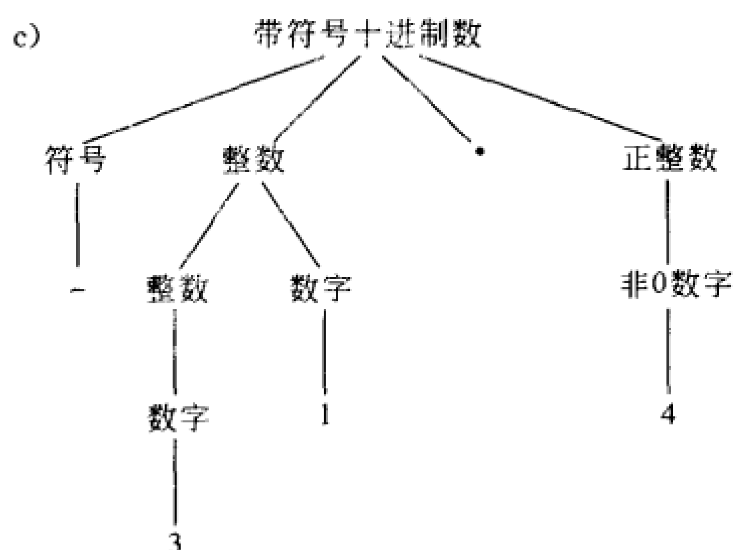
〈整数〉∷=〈数字〉|〈整数〉〈数字〉

〈数字〉∷=0|1|2|3|4|5|6|7|8|9

〈非 0 数字〉∷=1|2|3|4|5|6|7|8|9

〈正整数〉∷=〈整数〉〈非 0 数字〉|〈非 0 数字〉〈整数〉|〈整数〉〈非 0 数字〉〈整数〉|〈非 0 数字〉

c)



25. a)⟨identifier⟩ ::=⟨lcletter⟩ | ⟨identifier⟩⟨lcletter⟩

⟨lcletter⟩ ::=a | b | c | ... | z

b)⟨identifier⟩ ::=⟨lcletter⟩⟨lcletter⟩⟨lcletter⟩ | ⟨lcletter⟩⟨lcletter⟩⟨lcletter⟩⟨lcletter⟩ |
⟨lcletter⟩⟨lcletter⟩⟨lcletter⟩⟨lcletter⟩⟨lcletter⟩ |
⟨lcletter⟩⟨lcletter⟩⟨lcletter⟩⟨lcletter⟩⟨lcletter⟩⟨lcletter⟩

⟨lcletter⟩ ::=a | b | c | ... | z

c)⟨identifier⟩ ::=⟨ucletter⟩ | ⟨ucletter⟩⟨letter⟩ | ⟨ucletter⟩⟨letter⟩⟨letter⟩ |
⟨ucletter⟩⟨letter⟩⟨letter⟩⟨letter⟩ | ⟨ucletter⟩⟨letter⟩⟨letter⟩⟨letter⟩⟨letter⟩ |
⟨ucletter⟩⟨letter⟩⟨letter⟩⟨letter⟩⟨letter⟩⟨letter⟩

⟨letter⟩ ::=⟨lcletter⟩ | ⟨ucletter⟩

⟨lcletter⟩ ::=a | b | c | ... | z

⟨ucletter⟩ ::=A | B | C | ... | Z

d)⟨identifier⟩ ::=⟨lcletter⟩⟨digitorus⟩⟨alphanumeric⟩⟨alphanumeric⟩⟨alphanumeric⟩ |
⟨lcletter⟩⟨digitorus⟩⟨alphanumeric⟩⟨alphanumeric⟩⟨alphanumeric⟩⟨alphanumeric⟩

⟨digitorus⟩ ::=⟨digit⟩ | _

⟨alphanumeric⟩ ::=⟨letter⟩ | ⟨digit⟩

⟨letter⟩ ::=⟨lcletter⟩ | ⟨ucletter⟩

⟨lcletter⟩ ::=a | b | c | ... | z

⟨ucletter⟩ ::=A | B | C | ... | Z

⟨digit⟩ ::=0 | 1 | 2 | ... | 9

27. ⟨identifier⟩ ::=⟨letterorus⟩ | ⟨identifier⟩⟨symbol⟩

⟨letterorus⟩ ::=⟨letter⟩ | _

⟨symbol⟩ ::=⟨letterorus⟩ | ⟨digit⟩

⟨letter⟩ ::=⟨lcletter⟩ | ⟨ucletter⟩

⟨lcletter⟩ ::=a | b | c | ... | z

⟨ucletter⟩ ::=A | B | C | ... | Z

⟨digit⟩ ::=0 | 1 | 2 | ... | 9

29. numeral ::=sign? nonzerodigit digit * decimal? | sign? 0decimal?

sign ::=+ | −

nonzerodigit ::=1 | 2 | ... | 9

digit ::=0 | nonzerodigit

decimal ::=. digit *

31. identifier ::=letterorus symbol *

letterorus ::=letter | _

symbol ::=letterorus | digit

letter ::=lcletter | ucletter

lcletter ::=a | b | c | ... | z

ucletter ::=A | B | C | ... | Z

digit ::=0 | 1 | 2 | ... | 9

33. a)⟨expression⟩

⟨term⟩⟨term⟩⟨addOperator⟩

⟨factor⟩⟨factor⟩⟨factor⟩⟨mulOperator⟩⟨addOperator⟩

⟨identifier⟩⟨identifier⟩⟨identifier⟩⟨mulOperator⟩⟨addOperator⟩

abc * +

b)不可生成

c)⟨expression⟩

⟨term⟩

⟨factor⟩⟨factor⟩⟨mulOperator⟩

⟨expression⟩⟨factor⟩⟨mulOperator⟩

⟨term⟩⟨term⟩⟨addOperator⟩⟨factor⟩⟨mulOperator⟩

⟨factor⟩⟨factor⟩⟨addOperator⟩⟨factor⟩⟨mulOperator⟩

⟨identifier⟩⟨identifier⟩⟨addOperator⟩⟨identifier⟩⟨mulOperator⟩

x y — z *

d)⟨expression⟩

⟨term⟩

⟨factor⟩⟨factor⟩⟨mulOperator⟩

⟨factor⟩⟨expression⟩⟨mulOperator⟩

⟨factor⟩⟨term⟩⟨mulOperator⟩

⟨factor⟩⟨factor⟩⟨factor⟩⟨mulOperator⟩⟨mulOperator⟩

⟨factor⟩⟨factor⟩⟨expression⟩⟨mulOperator⟩⟨mulOperator⟩

⟨factor⟩⟨factor⟩⟨term⟩⟨term⟩⟨addOperator⟩⟨mulOperator⟩⟨mulOperator⟩

⟨factor⟩⟨factor⟩⟨factor⟩⟨factor⟩⟨addOperator⟩⟨mulOperator⟩⟨mulOperator⟩

⟨identifier⟩⟨identifier⟩⟨identifier⟩⟨identifier⟩⟨addOperator⟩⟨mulOperator⟩⟨mulOperator⟩

w x y z — * /

e)⟨expression⟩

⟨term⟩

⟨factor⟩⟨factor⟩⟨mulOperator⟩

⟨factor⟩⟨expression⟩⟨mulOperator⟩

⟨factor⟩⟨term⟩⟨term⟩⟨addOperator⟩⟨mulOperator⟩

⟨factor⟩⟨factor⟩⟨factor⟩⟨addOperator⟩⟨mulOperator⟩

⟨identifier⟩⟨identifier⟩⟨identifier⟩⟨addOperator⟩⟨mulOperator⟩

a d e — *

35. a)不可生成

b)⟨expression⟩

⟨term⟩⟨addOperator⟩⟨term⟩

⟨factor⟩⟨mulOperator⟩⟨factor⟩⟨addOperator⟩⟨factor⟩⟨mulOperator⟩⟨factor⟩

⟨identifier⟩⟨mulOperator⟩⟨identifier⟩⟨addOperator⟩⟨identifier⟩⟨mulOperator⟩⟨identifier⟩

a/b+c/d

c)⟨expression⟩

⟨term⟩

⟨factor⟩⟨mulOperator⟩⟨factor⟩

⟨factor⟩⟨mulOperator⟩(⟨expression⟩)

⟨factor⟩⟨mulOperator⟩(⟨term⟩⟨addOperator⟩⟨term⟩)

⟨factor⟩⟨mulOperator⟩(⟨factor⟩⟨addOperator⟩⟨factor⟩)

⟨identifier⟩⟨mulOperator⟩(⟨identifier⟩⟨addOperator⟩⟨identifier⟩)

m * (n+p)

d)不可生成

e) ⟨*expression*⟩

⟨*term*⟩

⟨*factor*⟩⟨*mulOperator*⟩⟨*factor*⟩

(⟨*expression*⟩)⟨*mulOperator*⟩(⟨*expression*⟩)

(⟨*term*⟩⟨*addOperator*⟩⟨*term*⟩)⟨*mulOperator*⟩(⟨*term*⟩⟨*addOperator*⟩⟨*term*⟩)

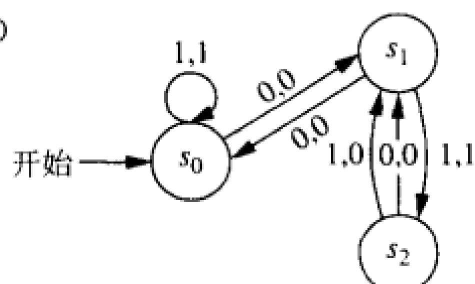
(⟨*factor*⟩⟨*addOperator*⟩⟨*factor*⟩)⟨*mulOperator*⟩(⟨*factor*⟩⟨*addOperator*⟩⟨*factor*⟩)

(⟨*identifier*⟩⟨*addOperator*⟩⟨*identifier*⟩)⟨*mulOperator*⟩(⟨*identifier*⟩⟨*addOperator*⟩⟨*identifier*⟩)

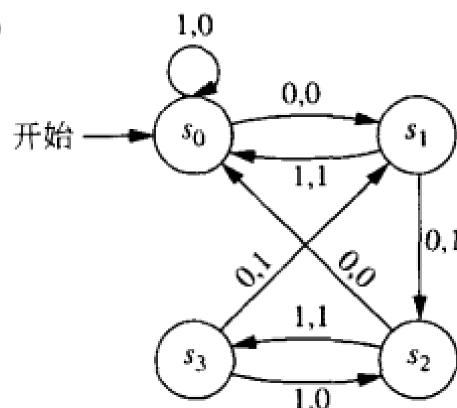
$(m+n)*(p-q)$

11.2 节

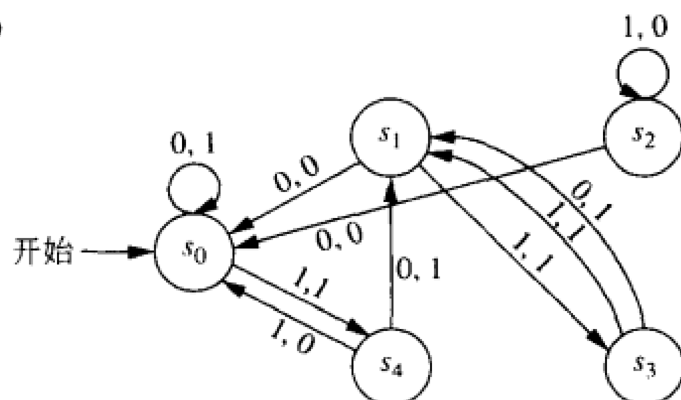
1. a)



b)

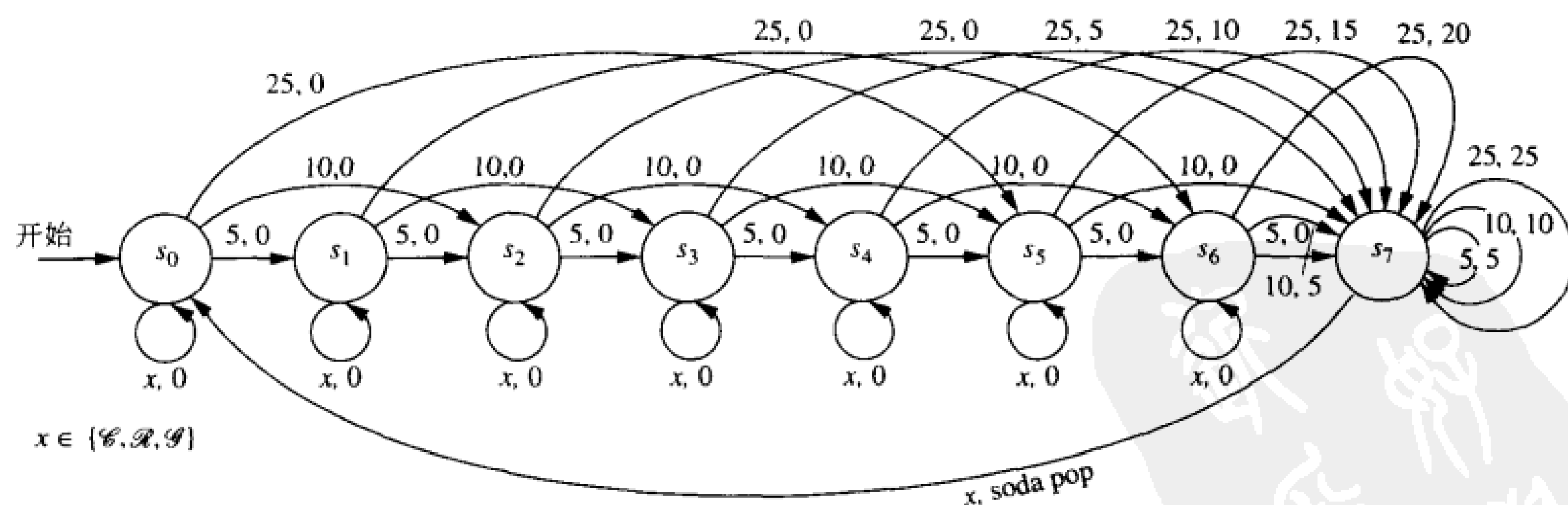


c)



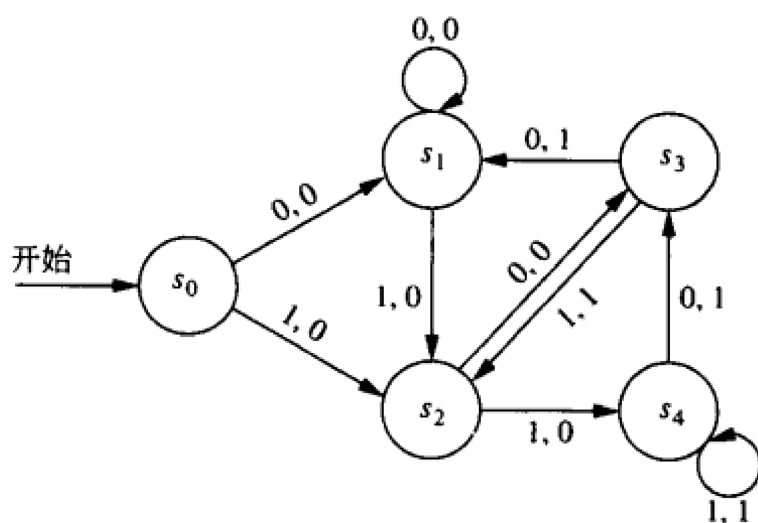
3. a) 1100　　b) 00110110　　c) 11111111111

5.

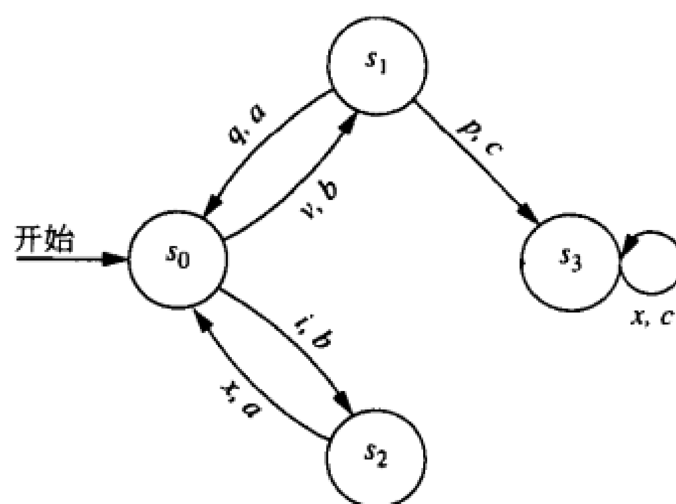


𝒞 = 可乐
ℛ = 无醇啤酒
𝒢 = 姜汁啤酒

7.

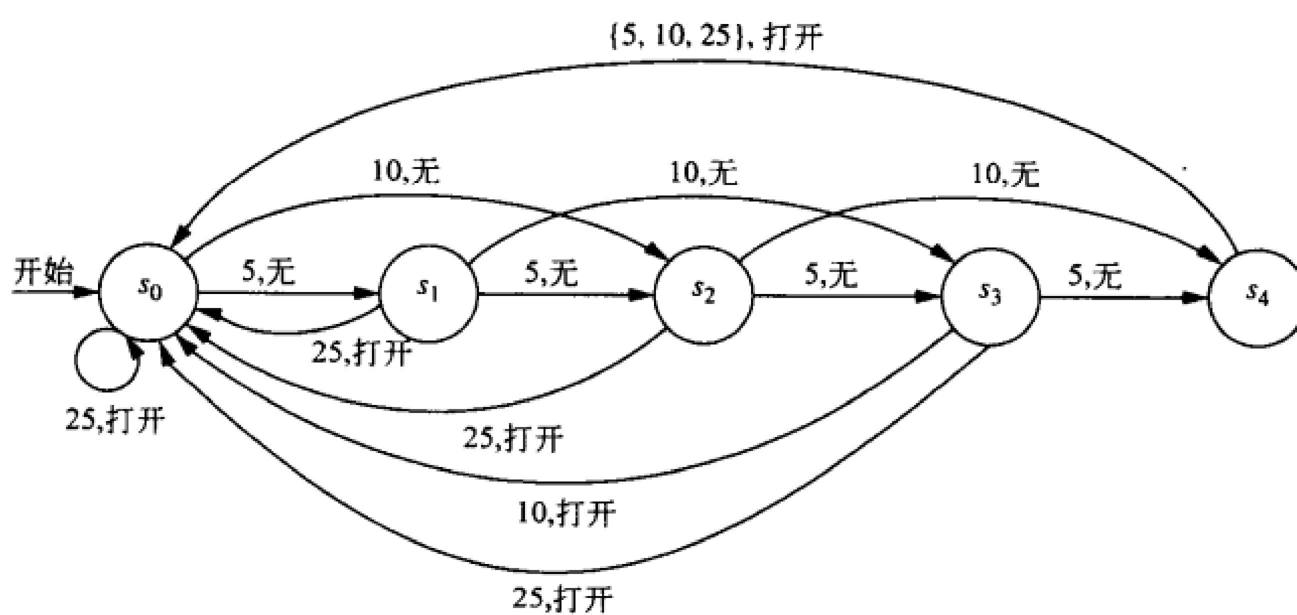


9.

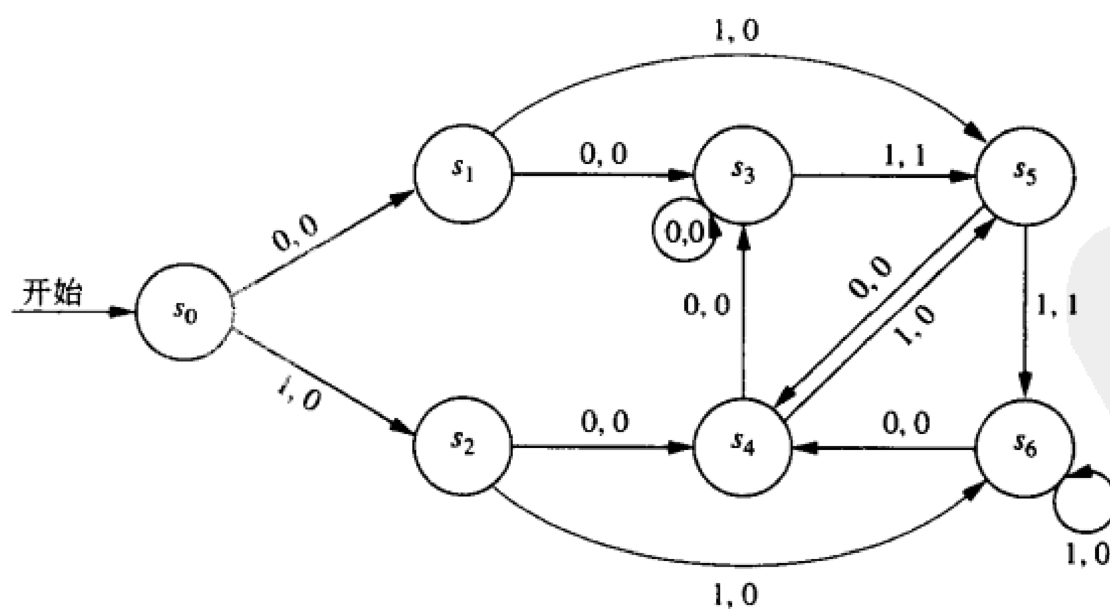


v = 有效标识	a = “输入用户标识”
i = 无效标识	b = “输入口令”
p = 有效口令	c = 提示
q = 无效口令	x = 任意输入

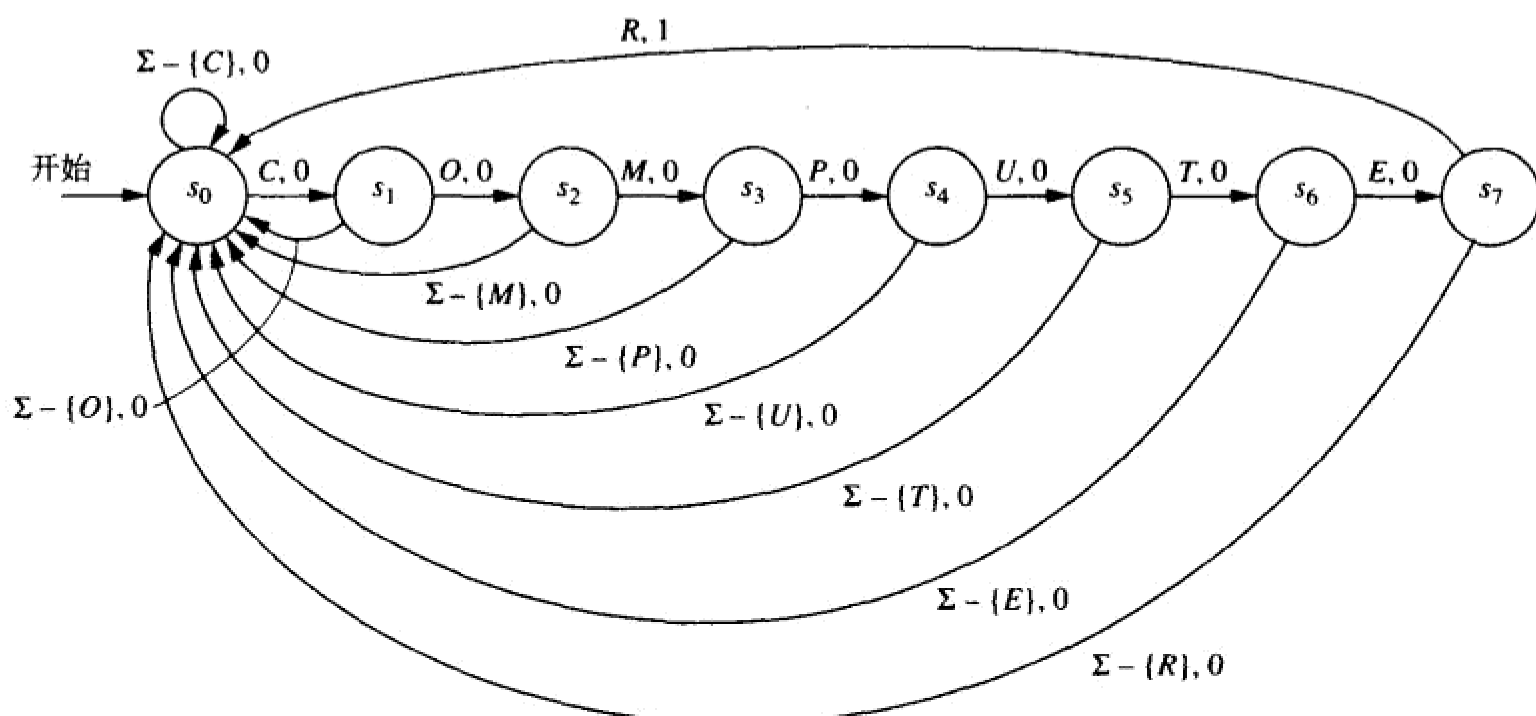
11.



13.



15.



17.

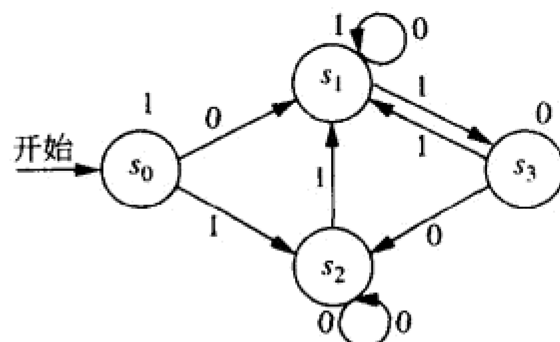
状态	f 输入 0	f 输入 1	g
s_0	s_1	s_2	1
s_1	s_1	s_0	1
s_2	s_1	s_2	0

19. a)11111

b)1000000

c)100011001100

21.



11.3节

1. a)$\{000, 001, 1100, 1101\}$　b)$\{000, 0011, 010, 0111\}$　c)$\{00, 011, 110, 1111\}$

 d)$\{000000, 000001, 000100, 000101, 010000, 010001, 010100, 010101\}$

3. $A=\{1, 101\}$, $B=\{0, 11, 000\}$; $A=\{10, 111, 1010, 1000, 10111, 101000\}$, $B=\{\lambda\}$;

 $A=\{\lambda, 10\}$, $B=\{10, 111, 1000\}$, 或者 $A=\{\lambda\}$, $B=\{10, 111, 1010, 1000, 10111, 101000\}$。

5. a)由零个或多个连续的二进制数对 10 构成的串。

 b)下列由 1 组成的串集合：串中 1 的个数能被 3 整除，包括空串。

 c)下列串的集合，以 0 开始，且每对 1 之间至少有一个 0。

 d)下列串的集合：以 1 开始和结尾，且每对 0 之间至少有两个 1。

7. 一个串在 A^* 中当且仅当它是 A 中任意多个串的连接。因为 A 的每个串都在 B 中，故 A^* 的每个串也是 B 中串的连接。从而 $A^* \subseteq B^*$。

9. a)是　b)是　c)是

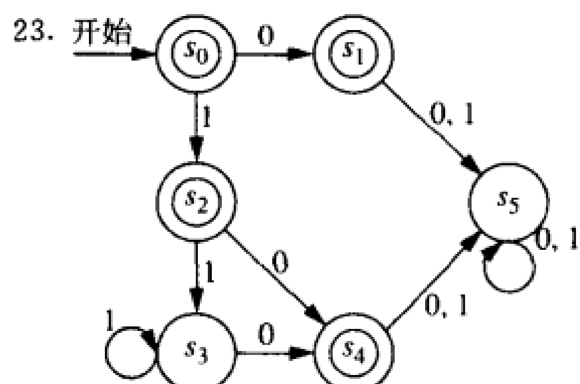
 d)不是　e)是　f)是

11. a)不是　b)是　c)不是

 d)不是　e)不是　f)不是

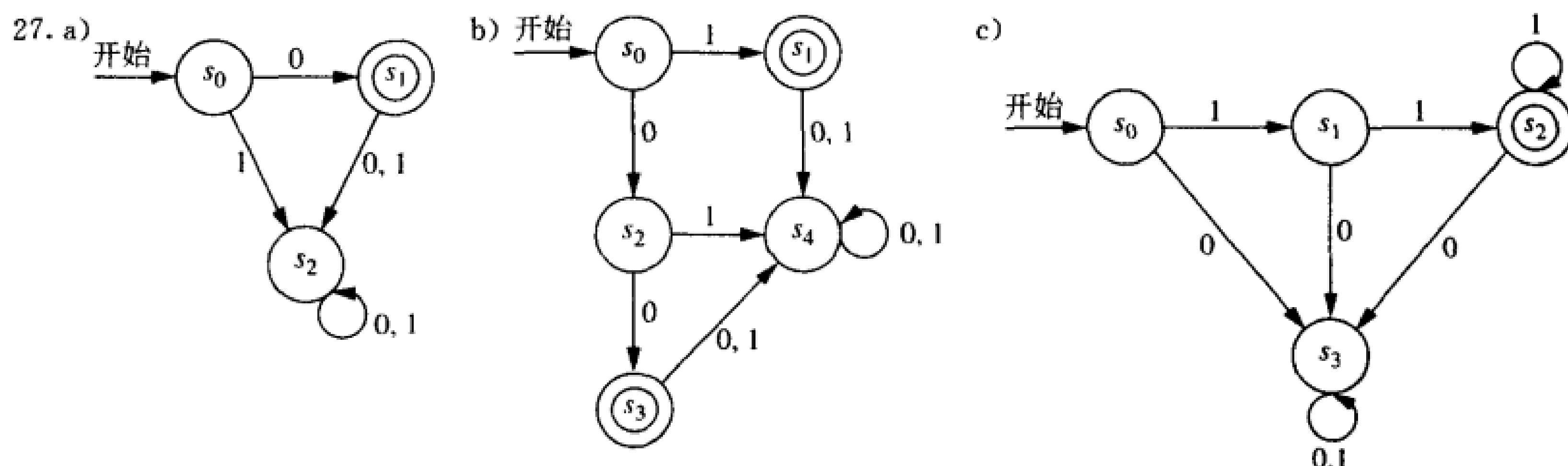
13. $\{0, 10, 11\}\{0, 1\}^*$　15. $\{0^m1^n \mid m \geqslant 0$ 且 $n \geqslant 1\}$　17. $\{0, 01, 11\}$

19. $\{\lambda, 0\} \cup \{0^m1^n \mid m \geqslant 1, n \geqslant 1\}$　21. $\{10^n \mid n \geqslant 0\} \cup \{10^n10^m \mid n, m \geqslant 0\}$

23. 

25. 增加非终结状态 s_3 和下列到 s_3 的转移：从 s_0 出发且输入 0 上、从 s_1 出发且在输入 1 上、从 s_3 出发且在输入 0 或 1 上。

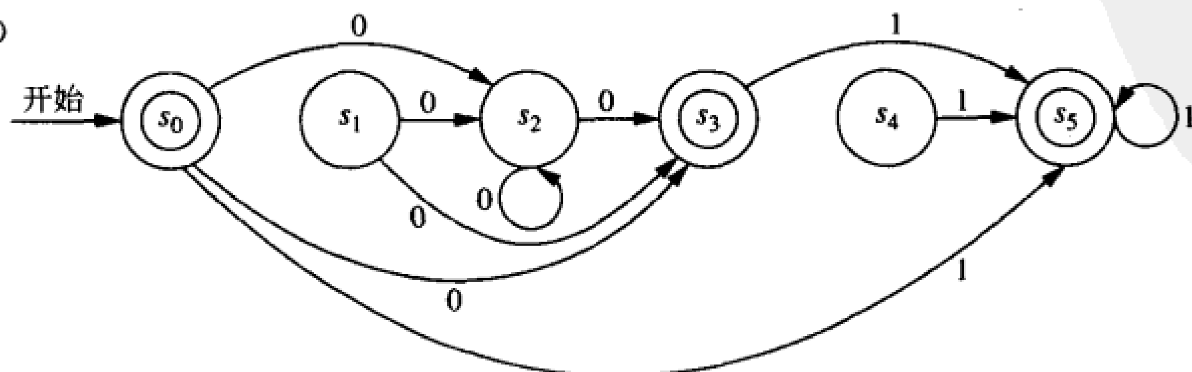
27. a) b) c)



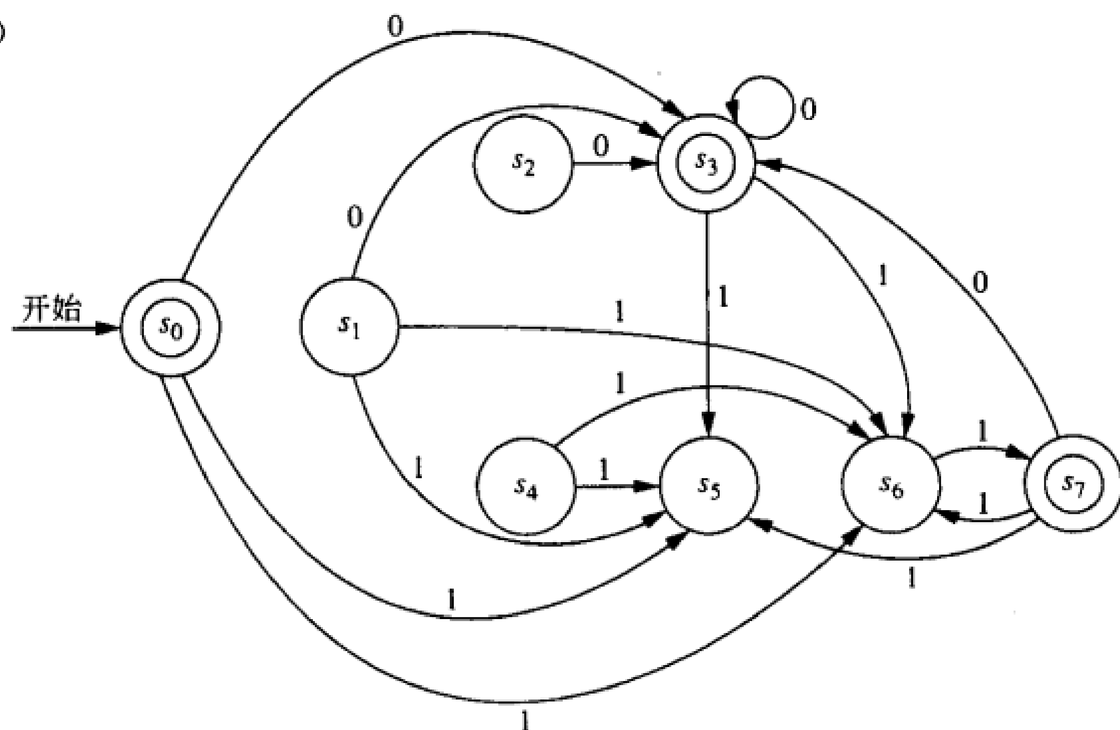
29. 设 M 是一个有限状态自动机，其接受的集合是含有相同 0 和 1 个数的二进制数串组成的集合。且设 M 有 n 个状态。考虑串 $0^{n+1}1^{n+1}$。由鸽巢原理知：当 M 处理这个串时，它在读前面 $n+1$ 个 0 的过程中必定不止一次遇到某个同样的状态。设 s 是一个它至少遇到两次的状态，则存在一个正整数 k 使得：输入中的 k 个 0 将 M 从状态 s 带回到自身。故 M 在读 $0^{n+1+k}1^{n+1}$ 之后的结束状态与读 $0^{n+1}1^{n+1}$ 之后的结束状态完全相同。因此，由 M 接受 $0^{n+1}1^{n+1}$ 知它也接受 $0^{n+1+k}1^{n+1}$，矛盾。

11.4 节

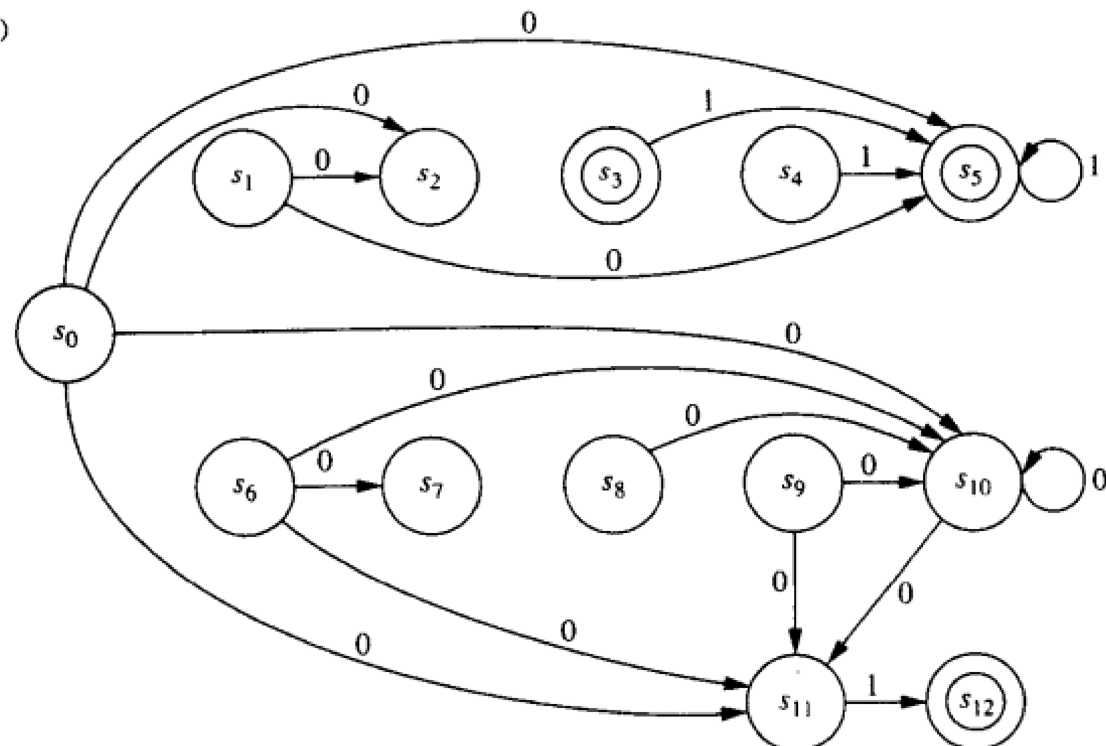
1. a)任意多个 1，后面跟个 0。
   b)任意多个 1，后面跟一个或一个以上的 0。
   c)111 或 011
   d)任意多个 1 组成的串，或任意多 00 组成的串，或者一列这样的串。
   e)λ，或者是以 1 结尾且在每个 1 之前至少有一个或多个 0 构成的串。
   f)长度至少为 3 且以 00 结束的串。
3. a)$\mathbf{00^*1}$　b)$\mathbf{(0\cup1)(0\cup1)(0\cup1)^*0000^*}$　c)$\mathbf{0^*1^*\cup1^*0^*}$　d)$\mathbf{11(111)^*(00)^*}$
5. 用归纳法证明。若 A 的正则表达式是 $\varnothing$、λ 或 x，结论显然成立。否则，设 A 的正则表达式是 $\mathbf{BC}$，则 $A=BC$，其中 B 是 $\mathbf{B}$ 生成的集合，C 是 $\mathbf{C}$ 生成的集合。由归纳假设，存在正则表达式 $\mathbf{B'}$ 和 $\mathbf{C'}$ 分别生成集合 B^R 和 C^R。因为 $A^R=(BC)^R=C^RB^R$，故 $\mathbf{C'B'}$ 是 A^R 的正则表达式。若 A 的正则表达式是 $\mathbf{B\cup C}$，则 $\mathbf{B'\cup C'}$ 是 A^R 的正则表达式，因为 $(B\cup C)^R=(B^R)\cup(C^R)$。最后，若 A 的正则表达式是 $\mathbf{B^*}$，易见：A^R 的正则表达式是 $\mathbf{(B')^*}$。
7. a)



b)



c)



9. $S\to0A$，$S\to1B$，$S\to0$，$A\to0B$，$A\to1B$，$B\to0B$，$B\to1B$。

11. $S\to0C$，$S\to1A$，$S\to1$，$A\to1A$，$A\to0C$，$A\to1$，$B\to0B$，$B\to1B$，$B\to0$，$B\to1$，$C\to0C$，$C\to1B$，$C\to1$。

13. 这是因为：此自动机中导致终结状态的输入对应于文法中唯一的产生式。

15. 因为 I 是有限的，故必要性显然成立。对于充分性，设状态为 s_{i_0}，s_{i_1}，s_{i_2}，…，s_{i_n}，其中 $n=l(x)$。因为 $n\geqslant|S|$，由鸽巢原理知，某个状态必定要重复。设 y 是 x 中引起循环的部分，即 $x=uyv$，其中 y 将某个 s_j 送到 s_j。则对任意 k，$uy^kv\in L(M)$。从而 $L(M)$是无限的。

17. 设 $L=\{0^{2n}1^n\mid n\geqslant0\}$是正则的，则有一个有限状态机器识别它，令 S 为此机器的状态集。设 $z=0^{2n}1^n$，其中 $3n\geqslant|S|$。则由泵引理，$z=0^{2n}1^n=uvw$，$l(v)\geqslant1$，且 $uv^iw\in\{0^{2n}1^n\mid n\geqslant0\}$。显然，$v$ 不能既包含 0 又包含 1，因为这样使得 v^2 包含 10。因此 v 全部都是 0 或全部都是 1，这样 uv^2w 包含了太多的 0 或太多的 1，因而它不在 L 中。此矛盾表明 L 不是正则的。

19. 假设$\{0,1\}$上的回文集是正则的，则有一个有限状态机器识别它，令 S 为此机器的状态集。设 $z=0^n10^n$，其中 $n>|S|$。由泵引理可找到 u、v 和 w 使 $uv^iw\in L$ 对所有非负整数 i 成立，其中，$l(v)\geqslant1$，$l(uv)\leqslant|S|$，且 $z=0^n10^n=uvw$。从而 v 必定是由 0 构成的串(因为 $|n|>|S|$)，故 uv^2w 不是回文。因此回文集不是正则的。

11.5节

1. a)当机器停机时，带的非空白区域包含串1111。
   b)当机器停机时，带的非空白区域包含串011。
   c)当机器停机时，带的非空白区域包含串00001。
   d)当机器停机时，带的非空白区域包含串00。
3. 如果带上至少有一个1，则机器从第一个1开始，每隔一个1变为0，且在到达第一个空白符时停机。若一开始的时候带就是空白，则机器不对带作任何改变就停机。如果带的非空白区域所含的都是0，则机器在相继穿过这些0后停机。
5. $(s_0, 0, s_1, 1, R)$，$(s_0, 1, s_0, 1, R)$
7. $(s_0, 0, s_0, 0, R)(s_0, 1, s_1, 1, R)$，$(s_1, 0, s_1, 0, R)$，$(s_1, 1, s_1, 0, R)$。
9. $(s_0, 0, s_1, 0, R)$，$(s_0, 1, s_0, 0, R)$，$(s_1, 0, s_1, 0, R)$，$(s_1, 1, s_0, 0, R)$，(s_1, B, s_2, B, R)。
11. $(s_0, 0, s_0, 0, R)$，$(s_0, 1, s_1, 1, R)$，$(s_1, 0, s_1, 0, R)$，$(s_1, 1, s_0, 1, R)$，(s_0, B, s_2, B, R)。
13. 若输入串为空串或以1开头，则机器在非终结状态S_0下停机。若不是，则开头的0变为M，且机器掠过中间所有的0和1，直到到达输入串的尾部或遇到M，此处它往后退一个方格，并进入状态s_2。因为可接受串对于左边的每个0都必须在右边有个1，故若此串是可接受的，此处必定有个1。当这个方格所含的是1时，从s_2出发的唯一转移就要发生。如果这样，机器就用M取代之，且向左继续回走；如果不是这样，则机器在非终结状态s_2下停机。在往回走的过程中，如果它所见的是1，则处于s_3；如果它所见的是0，则处于s_4。最后它处于下列两种情形之一：一是在s_4下遇到1，此时它停机但不接受；或者到达已经在此串初始部分的0上写下的最右的M。若此发生时它处于s_3，则此串中不再有其他的0，因此最好也没有其他的1。这可以用转移(s_3, M, s_5, M, R)和(s_5, M, s_6, M, R)来完成，其中s_6是一个终结状态。否则机器在非终结状态s_5下停机。若遇到这个M时处于s_4，则事情又完全从头开始，除了将串中剩下最左边的0和最右边的1用M来替代。所以机器在状态s_4下移动到剩下最左边的0，且回到状态s_0再重复这个过程。
15. (s_0, B, s_9, B, L)，$(s_0, 0, s_1, 0, L)$，(s_1, B, s_2, E, R)，(s_2, M, s_2, M, R)，$(s_2, 0, s_3, M, R)$，$(s_3, 0, s_3, 0, R)$，(s_3, M, s_3, M, R)，$(s_3, 1, s_4, M, R)$，$(s_4, 1, s_4, 1, R)$，(s_4, M, s_4, M, R)，$(s_4, 2, s_5, M, R)$，$(s_5, 2, s_5, 2, R)$，(s_5, B, s_6, B, L)，(s_6, M, s_8, M, L)，$(s_6, 2, s_7, 2, L)$，$(s_7, 0, s_7, 0, L)$，$(s_7, 1, s_7, 1, L)(s_7, 2, s_7, 2, L)$，$(s_7, M, s_7, M, L)$，$(s_7, E, s_2, E, R)$，$(s_8, M, s_8, M, L)$，$(s_8, E, s_9, E, L)$，其中$M$和$E$是记号，$E$标记输入的左端点。
17. $(s_0, 1, s_1, B, R)$，$(s_1, 1, s_2, B, R)$，$(s_2, 1, s_3, B, R)$，$(s_3, 1, s_4, 1, R)$，$(s_1, B, s_4, 1, R)$，$(s_2, B, s_4, 1, R)$，$(s_3, B, s_4, 1, R)$。
19. $(s_0, 1, s_1, B, R)$，$(s_1, 1, s_2, B, R)$，(s_1, B, s_6, B, R)，$(s_2, 1, s_3, B, R)$，(s_2, B, s_6, B, R)，$(s_3, 1, s_4, B, R)$，(s_3, B, s_6, B, R)，$(s_4, 1, s_5, B, R)$，(s_4, B, s_6, B, R)，$(s_6, B, s_{10}, 1, R)$，$(s_5, 1, s_5, B, R)$，$(s_5, B, s_7, 1, R)$，$(s_7, B, s_8, 1, R)$，$(S_8, B, s_9, 1, R)$，$(s_9, B, s_{10}, 1, R)$。
21. $(s_0, 0, s_0, 0, R)$，$(s_0, *, s_5, B, R)$，$(s_3, *, s_3, *, L)$，$(s_3, 0, s_3, 0, L)$，$(s_3, 1, s_3, 1, L)$，(s_3, B, s_0, B, R)，$(s_5, 1, s_5, B, R)$，$(s_5, 0, s_5, B, R)$，(s_5, B, s_6, B, L)，(s_6, B, s_6, B, L)，$(s_6, 0, s_7, 1, L)$，$(s_7, 0, s_7, 1, L)$，$(s_0, 1, s_1, 0, R)$，$(s_1, 1, s_1, 1, R)$，$(s_1, *, s_2, *, R)$，$(s_2, 0, s_2, 0, R)$，$(s_2, 1, s_3, 0, L)$，(s_2, B, s_4, B, L)，$(s_4, 0, s_4, 1, L)$，$(s_4, *, s_8, B, L)$，$(s_8, 0, s_8, B, L)$，$(s_8, 1, s_8, B, L)$。
23. $(s_0, B, s_1, 1, L)$，$(s_0, 1, s_1, 1, R)$，$(s_1, B, s_0, 1, R)$。

补充练习

1. a)$S\to 00S111$，$S\to\lambda$。
   b)$S\to AABS$，$AB\to BA$，$BA\to AB$，$A\to 0$，$B\to 1$，$S\to\lambda$。
   c)$S\to ET$，$T\to 0TA$，$T\to 1TB$，$T\to\lambda$，$0A\to A0$，$1A\to A1$，$0B\to B0$，$1B\to B1$，$EA\to E0$，$EB\to E1$，$E\to\lambda$。

3.

5.

7. 不是。取 $A=\{1, 10\}$，$B=\{0, 00\}$。

9. 不是。取 $A=\{00, 000, 00000\}$，$B=\{00, 000\}$。

11. a)1 b)1 c)2 d)3 e)2 f)4

13.

15.

17. a) $n^{nk+1}m^{nk}$ b) $n^{nk+1}m^{n}$

19.

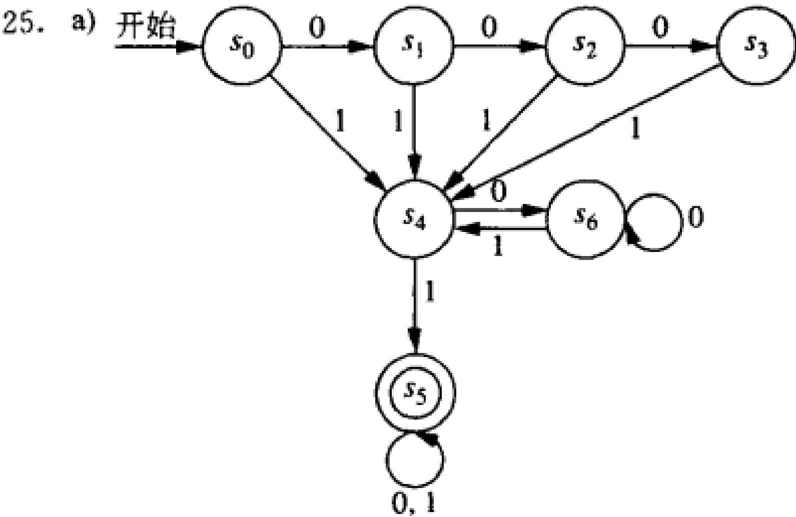
21. a)

b)

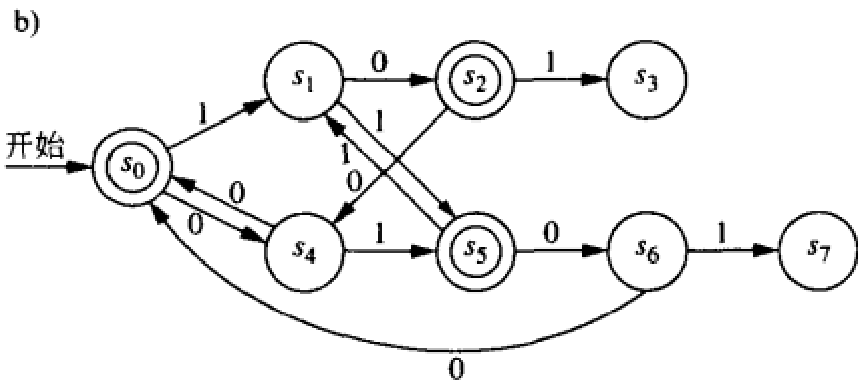
c)

23. 用状态 S 和终结状态 F 构造 A 的确定型有限状态自动机。对 $\bar{A}$ 使用相同的自动机，但其终结状态为 $S-F$。

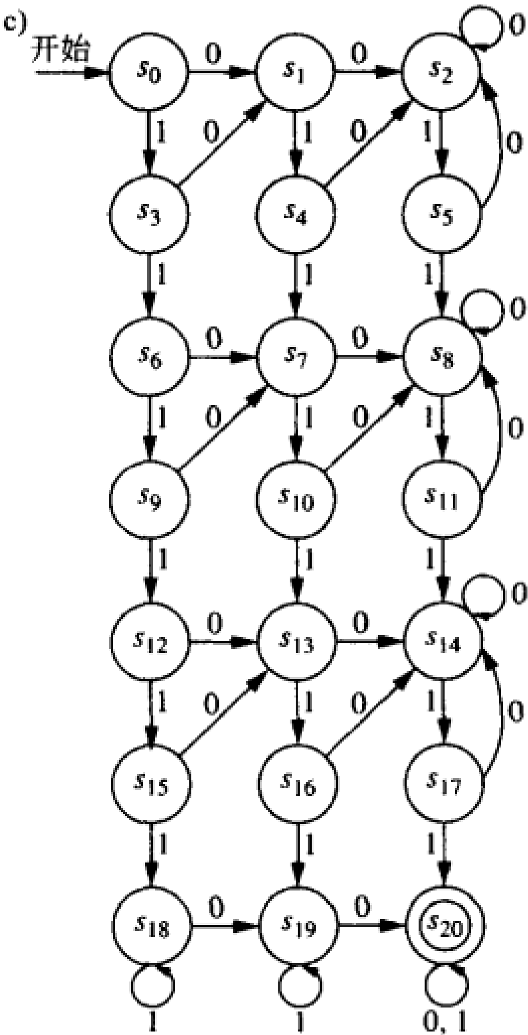
25. a)



b)



c)



27. 假设 $L=\{1^p \mid p$ 是素数$\}$是正则的，则有一个有限状态机器识别它，令 S 为此机器的状态集。设 $z=1^p$，其中 p 是一个大于 $|S|$ 的素数(这样的素数是存在的，因为有无限多个素数)。由泵引理知：必定能够将 z 写成 $z=uvw$，且 $l(uv)\leqslant|S|$，$l(v)\geqslant 1$。并且对任意非负整数 i，$uv^iw\in L$。因为 z 是由 1 构成的串，可令 $u=1^a$，$v=1^b$，$w=1^c$，其中 $a+b+c=p$，$a+b\leqslant n$，且 $b\geqslant 1$。这意味着 $uv^iw=1^a1^{bi}1^c=1^{(a+b+c)+b(i-1)}=1^{p+b(i-1)}$。现在取 $i=p+1$，则 $uv^iw=1^{p(1+b)}$。因为 $p(1+b)$不是素数，故 $uv^iw\notin L$，矛盾。

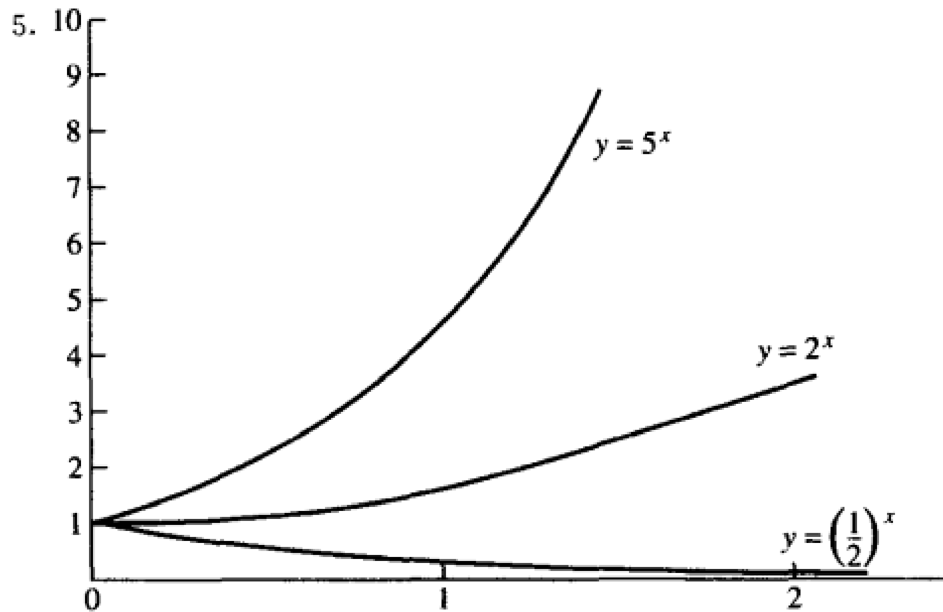
附录

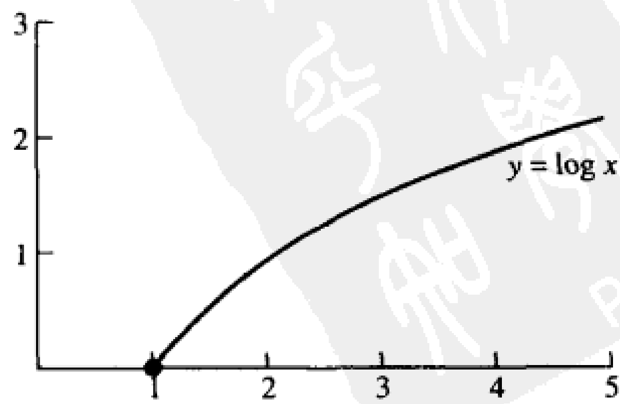
附录 A

1. a) 2^3　　b) 2^6　　c) 2^4

3. a) $2y$　　b) $2y/3$　　c) $y/2$

5.





附录 B

1. 第一块执行后，b 原来的值被赋给 a，c 原来的值被赋给 b。但第二块执行之后，c 原来的值被赋给 b，且 c 原来的值也被赋给了 a。
3. 下面的 **while** 结构实现相同的事情：

```
i := 起始值
while i ≤ 结束值
begin
    语句
    i := i+1
end
```



推荐读物

提出的书面材料和相关的网站都是用来进一步学习本书内容的资源。本推荐读物描述的书面材料，将分章列出，并与特定研究课题相联系。有一些内容全面的参考书特别值得一提。读者可能发现，Rosen 的《离散和组合数学手册》(*Handbook of Discrete and Combinatorial Mathematics*)[Ro00a]非常有用，这是一本综合参考书。在 Michaels 和 Rosen[MiRo91]中，可以找到离散数学的其他应用。在 Gruska[Gr97]中，有计算机科学中许多课题(包括本书中讨论的课题)的更深入介绍。本书提到的许多有关数学和计算机科学的文献信息可以在 Gillispie[Gi70]中查到。

要查找相关的网站，请在本书的网站中查询。本书网站的统一资源定位器(URL)是 www. mhhe. com/rosen。

第 1 章

学习逻辑的一个有趣方法是读 Lewis Carroll[Ca78]一书。逻辑的一般参考书包括 Mendelson[Me64]、Stoll[St74]和 Suppes[Su87]。Gries 和 Schneider[GrSc93]中可以找到离散数学中的逻辑学的全面介绍。系统规范在 Ince[In93]中有讨论。Smullyan 的骑士和流氓的难题在[Sm78]中介绍，他还写了许多有关逻辑难题的书，包括[Sm92]和[Sm98]。Nilsson 和 Maluszynski[NiMa95]中有 Prolog 的深入介绍，Clocksin 和 Mellish[ClMe94]中也有。Cupillari[Cu01]、Solow[So01]、Velleman[Ve94]及 Wolf[Wo98]中介绍了证明方法的基础。Lin 和 Lin[LiLi81]是关于集合论及其应用的一本易读的教材。公理集合论的进展可在 Halmos[Ha60]、Monk[Mo69]或 Stoll[St74]中找到。Brualdi[Br77]或 Reingold、Nievergelt 和 Deo[ReNiDe77]中有关于多重集合的介绍。Negoita[Ne85]和 Zimmerman[Zi91]都论述了模糊集合及其在专家系统和人工智能中的应用。微积分的书中，如 Apostol[Ap67]、Spivak[Sp94]及 Thomas 与 Finney[ThFi96]，都有关于函数的讨论。在 Arbib、Kfoury 与 Moll[ArKfMo80]，Bobrow 与 Arbib[BoAr74]，Beckman[Be80]和 Tremblay 与 Manohar[TrMa75]中，都可找到有关计算机科学的数学基础的讨论。本书提及了许多数学家和计算机科学家，他们的传记可在 Gillispie[Gi70]中找到。

第 2 章

Knuth 的文章[Kn77]与 Wirth 的文章[Wi84]是算法这门学科的入门文献。在 Knuth[Kn97a]中可以找到函数大 O 估计的广泛讨论。有关算法及其复杂性的一般参考书有 Aho、Hopcroft 和 Ullman[AhHoUl74]，Baase 和 Van Gelder[BaGe99]，Cormen、Leierson、Rivest 和 Stein[CoLeRist01]，Gonnet[Go84]，Goodman 和 Hedetniemi[GoHe77]，Harel[Ha87]，Horowitz 和 Sahni[HoSa82]，Knuth 关于计算机程序设计艺术的著名系列著作[Kn97a，b，98]，Kronsjö[Kr87]，Pohl 和 Shaw[PoSh81]，Purdom 和 Brown[PuBr85]，Rawlins[Ra92]，Sedgewick[Se88]，Wilf[Wi86]以及 Wirth[Wi76]。Knuth[98]中有排序和搜索算法及其复杂性的详细介绍。有关数论的参考书有 Hardy 和 Wright[HaWr79]，Leveque[Le77]，Rosen[Ro00]以及 Stark[St78]。介绍数论在密码学中应用的书有 Denning[De82]，Menezes、van Oorschot 和 Vanstone[MeOoVa97]，Rosen[Ro00]，Seberry 和 Pieprzyk[SePi89]，Sinkov[Si66]以及 Stinson[St02]。Rivest、Shamir 和 Adleman[RiShAd78]介绍了 RSA 公钥系统。Knuth[Kn97b]以及 Pohl 与 Shaw 的[PoSh81]中讨论了有关计算机算术的算法。所有的线性代数书中都有矩阵及其运算的介绍，如 Curtis[Cu84]和 Strang[St88]。

第3章

Pólya 的三本书[Po62、Po71、Po90]以令人愉快的方式讨论了构造证明的科学与艺术。有关整数序列的最佳书面材料是 Sloan 和 Plouffe[SlPl95]。Stanat 和 McAllister[StMc77]中有一整节讨论可数性。在 Sominskii[So61]的英语译本(原文为俄文)中可以找到数学归纳法的通俗介绍。Golomb[Go94]中讨论了用 L 形片或其他类型的片来覆盖一个棋盘中的部分或全部方格问题。包含数学归纳法和递归定义全面介绍的书有 Liu[Li85],Sahni[Sa85],Stanat 和 McAllister[StMc77]以及 Tremblay 和 Manohar 的[TrMa75]。阿克曼函数是由阿克曼(Ackermann)在 1922 年引入的,它起源于递归函数论(例如可参见 Beckman[Be80]和 McNaughton[Mc82])和某些集合论算法的复杂性分析(见 Tarjan[Ta83])。Roberts[Ro86]、Rohl[Ro84]和 Wand[Wa80]都讨论了递归。程序正确性的讨论以及用来证明程序是正确的逻辑机制的讨论可在 Alagic 和 Arbib[AlAr78]、Anderson[An79]、Backhouse[Ba86]、Sahni[Sa85]以及 Stanat 和 McAllister[StMc77]中找到。

第4章

计数技术及其应用的一般参考书有 Anderson[An74]、Berman 和 Fryer[BeFr72]、Bogart[Bo86]、Bose 和 Manvel[BoMa86]、Brualdi[Br97]、Cohen[Co78]、Grimaldi[Gr94]、Liu[Li68]、Pólya、Tarjan 和 Woods[PoTaWo83],Riordan[Ri58]、Roberts[Ro84]、Tucker[Tu02]以及 Williamson[Wi85]。Vilenkin[Vi71]中有选择地介绍了一些组合问题及它们的解,在 Lovász[Lo79]中可以找到一些更困难的组合问题。关于因特网协议寻址和数据报的信息可在 Comer[Co00]中找到。鸽巢原理的应用可在 Brualdi[Br77]、Liu[Li85]和 Roberts[Ro84]中找到。在 Riordan[Ri68]中可以找到许多组合数学中的恒等式。Even[Ev73]、Lehmer[Le64]以及 Reingold、Nievergelt 和 Deo[ReNiDe77]都描述了组合算法(包括生成置换和组合的算法)。

第5章

离散概率论的参考书有 Feller[Fe68]、Nabin[Na00]和 Ross[Ro01]。Ross[Ro02]关注于概率论在计算机科学中的应用。Aho 和 Ullman[AhUl95]包括了计算机科学中与概率论相关的讨论,包括概率的编程应用。Aigner 和 Ziegler[AiZi01]中专门有一章讨论概率方法。Alon 和 Spencer[AlSp00]中更加全面地讨论了概率方法。

第6章

在 Roberts[Ro84]和 Tucker[Tu02]中可以找到使用递推关系的许多不同模型。在 Brualdi[Br77]、Liu[Li68]和 Mattson[Ma93]中可以找到常系数线性齐次递推关系以及相关的非齐次递推关系的详尽介绍。Roberts[Ro84]及 Stanat 和 McAllister[StMc77]都介绍了"分而治之"算法及其复杂性。Aho、Hopcroft 和 Ullman[AhHoUl74]及 Knuth[Kn81]描述了更快的整数乘法和矩阵乘法。Pólya、Tarjan 和 Woods[PoTaWo83]中有生成函数的极好介绍。详细研究生成函数的文献有 Brualdi[Br77]、Cohen[Co78]、Graham、Knuth 和 Patashnik[GrKnPa94]、Grimaldi[Gr94]和 Roberts[Ro84]。容斥原理的其他应用可在 Liu[Li85,Li68]、Roberts[Ro84]和 Ryser[Ry63]中找到。

第7章

关于关系(包括等价关系和偏序)的通用参考书有 Bobrow 和 Arbib[BoAr74]、Grimaldi[Gr94]、Sanhi[Sa85]以及 Tremblay 和 Manohar[TrMa75]。Date[Da82]及 Aho 和 Ullman[AhUl95]讨论了数据库的关系模型。Roy 和 Warshall 求传递闭包的原始文献分别在[Ro59]和[Wa62]中。研究有向图的文献有 Chartrand 和 Lesniak[ChLe86]、Gross 和 Yellen[GrYe99]、Robinson 和 Foulds[RoFo86]、Roberts[Ro84]以及 Tucker[Tu85]。Denning[De82]讨论了格在信

息流处理中的应用。

第 8 章

图论的一般参考书有 Behzad 和 Chartrand[BeCh71]、Chartrand 和 Lesniak[ChLe96]、Bondy 和 Murty[BoMu76]、Chartrand 和 Oellermann[ChOe93]、Graver 和 Watkins[GrWa77]、Gross 和 Yellen[GrYe99]、Harary[Ha69]、Ore[Or63]、Roberts[Ro84]、Tucker[Tu85]、Wilson[Wi85] 以及 Wilson 和 Watkins[WiWa90]。介绍图论广泛应用文献的有 Chartrand[Ch77]、Deo[De74]、Foulds[Fo92]、Roberts[Ro84、Ro76]、Wilson 和 Beineke[WiBe79]以及 McHugh[Mc90]。涉及大图(包括 Web 图)的应用在 Hayes[Ha00a, 00b]中讨论。Gibbons[Gi85]全面描述了图论算法。图论算法的其他参考书还有 Buckley 和 Harary[BuHa90], Chartrand 和 Oellermann[ChOe93]、Chachra、Ghare 和 Moore[ChGhMo79]、Even[Ev73、Ev79]、Hu[Hu82]以及 Reingold、Nievergelt 和 Deo[ReNiDe77]。[Eu53]是欧拉关于哥尼斯堡桥问题原文的译文。Gibbons[Gi85]、Liu[Li85]以及 Reingold、Nievergelt 和 Deo[ReNiDe77]都讨论了 Dijkstra 算法，Dijkstra 的原文在[Di59]中。Kuratwoski 定理的证明可在 Harary[Ha69]和 Liu[Li68]中找到。Chartrand 和 Lesniak [ChLe96]研究了图的交叉数和厚度。有关图着色及四色定理的参考书有 Barnette[Ba83]及 Saaty 和 Kainen[SaKa86]。Appel 和 Haken[ApHa76]报道了四色定理的最初结果。Roberts[Ro84]描述了图着色的应用。Biggs、Lloyd 和 Wilson[BiLlWi86]介绍图论的历史。Akl[Ak89]与 Siegel 和 Hsu[SiHs88]讨论了并行处理的互连网络。

第 9 章

下列文献都讨论了树：Deo[De74]、Grimaldi[Gr94]、Knuth[Kn97a]、Roberts[Ro84]、Tucker[Tu85]。Gotlieb 和 Gotlieb[GoGo78]、Horowitz 和 Sahni[HoSa82]以及 Knuth[Kn97a, 98]描述了树在计算机科学中的应用。Roberts[Ro84]介绍了树在许多不同领域中的应用。Hamming[Ha80]介绍了前缀码和霍夫曼码。回溯法是一种旧方法，它在解迷宫难题中的应用可从[Lu91]中找到，这是 Lucas 于 1891 年写的一本书。Reingold、Nievergelt 和 Deo[ReNiDe77]广泛讨论了怎么用回溯法来求解问题。Gibbons[Gi85]和 Reingold、Nievergelt 和 Deo[ReNiDe77]还讨论了构造生成树和极小生成树算法。Graham 和 Hell[GrHe85]介绍了求极小生成树算法的历史和背景。Prim 和 Kruskal 分别在[Pr57]和[Kr56]中描述了他们的求极小生成树算法。Sollin 算法是一个非常适用于并行处理算法的例子。虽然 Sollin 从未发表这个算法的描述，但 Even[Ev73]及 Goodman 和 Hedetniemi[GoHe77]都描述了它。

第 10 章

讨论布尔代数的文献有 Hohn[Ho66]、Kohavi[Ko86]以及 Tremblay 和 Manohar[TrMa75]。Hayes[Ha93]、Hohn[Ho66]、Katz[Ka93]和 Kohavi[Ko86]描述了布尔代数在逻辑电路和开关电路中的应用。用图来处理乘积的和展开式的极小化的原始文献是 Karnaugh[Ka53]和 Veitch [Ve52]。奎因-莫可拉斯基方法是在 McCluskey[Mc56]和 Quine[Qu52, Qu55]中引入的。Kohavi [Ko78]介绍了阈值函数。

第 11 章

语言和自动机理论的一般参考书有 Denning、Dennis 和 Qualitz[DeDeQu81]、Hopcroft 和 Ullman[HoUl79]、Hopkin 和 Moss[HoMo76]、Lewis 和 Papadimitriou[LePa84]以及 McNaughton[Mc82]。米利机和摩尔机是在 Mealy[Me55]和 Moore[Mo56]中引入的。克莱因定理的原始证明可在[Kl56]中找到。下列文献讨论了更强大的计算模型(包括下推自动机和图灵机)：Brookshear[Br89]、Hennie[He77]、Hopcroft 和 Ullman[HoUl79]、Hopkin 和 Moss[HoMo76]、Martin[Ma91]以及 Wood[Wo87]。Barwise 和 Etchemendy[BaEt93]是关于图灵机的一个不错的介

绍。在 Herken[He88]中可以找到一些关于图灵机及相关机器的历史和应用的有趣文章。忙碌海狸机器首先由 Rado 在[Ra62]中引入，有关这些机器的信息可在下列文献中找到：Dewdney[De84，93]、Herken[He88]中由 Brady 所写的文章以及 Wood[Wo87]。

附录

指数和对数函数的详细介绍可在微积分书中找到。如 Apostol[Ap67]、Spivak[Sp80]及 Thomas 和 Finney[ThFi92]。附录 2 中描述的伪代码与 Pascal 十分相像。Pohl 和 Shaw[PoSh81]用类似形式语言描述算法。Wirth[Wi76]描述了怎么用算法和数据结构来构造 Pascal 程序。Rohl[Ro84]说明怎么用 Pascal 来实现递归程序。



参考文献

[AhHoUl74] A. V. Aho, J. E. Hopcroft, and J. D. Ullman, *The Design and Analysis of Computer Algorithms,* Addison-Wesley, Reading, MA, 1974.

[AhUl95] Alfred V. Aho and Jeffrey D. Ullman, *Foundations of Computer Science,* C ed., Computer Science Press, New York, 1995.

[AiZi01] Martin Aigner and Günter M. Ziegler, *Proofs from THE BOOK,* 2d ed., Springer, Berlin, 2001.

[Ak89] S. G. Akl, *The Design and Analysis of Parallel Algorithms,* Prentice Hall, Englewood Cliffs, NJ, 1989.

[AlAr78] S. Alagic and M. A. Arbib, *The Design of Well-Structured and Correct Programs,* Springer-Verlag, New York, 1978.

[AlSp00] Noga Alon and Joel H. Spencer, *The Probabilistic Method,* 2d ed., Wiley, New York, 2000.

[An74] I. Anderson, *A First Course in Combinatorial Mathematics,* Clarendon, Oxford, England, 1974.

[An79] R. B. Anderson, *Proving Programs Correct,* Wiley, New York, 1979.

[Ap67] T. M. Apostol, *Calculus,* vol. I, 2d ed., Wiley, New York, 1967.

[ApHa76] K. Appel and W. Haken, "Every Planar Map Is 4-colorable," *Bulletin of the AMS,* 82 (1976), 711–712.

[ArKfMo80] M. A. Arbib, A. J. Kfoury, and R. N. Moll, *A Basis for Theoretical Computer Science,* Springer-Verlag, New York, 1980.

[BaGe99] S. Baase and A. Van Gelder, *Computer Algorithms: Introduction to Design and Analysis,* 3d ed., Addison-Wesley, Reading, MA, 1999.

[Ba86] R. C. Backhouse, *Program Construction and Verification,* Prentice-Hall, Englewood Cliffs, NJ, 1986.

[Ba83] D. Barnette, *Map Coloring, Polyhedra, and the Four-Color Problem,* Mathematical Association of America, Washington, D.C., 1983.

[BaEt93] Jon Barwise and John Etchemendy, *Turing's World 3.0 for the Macintosh,* CSLI Publications, Stanford, CA, 1993.

[Be80] F. S. Beckman, *Mathematical Foundations of Programming,* Addison-Wesley, Reading, MA, 1980.

[BeCh71] M. Behzad and G. Chartrand, *Introduction to the Theory of Graphs,* Allyn & Bacon, Boston, 1971.

[Be86] J. Bentley, *Programming Pearls,* Addison-Wesley, Reading, MA, 1986.

[BeFr72] G. Berman and K. D. Fryer, *Introduction to Combinatorics,* Academic Press, New York, 1972.

[BiLlWi86] N. L. Biggs, E. K. Lloyd, and R. J. Wilson, *Graph Theory 1736–1936,* Clarendon, Oxford, England, 1986.

[BoAr74] L. S. Bobrow and M. A. Arbib, *Discrete Mathematics,* Saunders, Philadelphia, 1974.

[Bo86] K. P. Bogart, *Introductory Combinatorics,* Wiley, New York, 1986.

[BoMu76] J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications,* North-Holland, New York, 1976.

[BoMa86] R. C. Bose and B. Manvel, *Introduction to Combinatorial Theory,* Wiley, New York, 1986.

[Br89] J. G. Brookshear, *Theory of Computation,* Benjamin Cummings, Redwood City, CA, 1989.

[Br97] R. A. Brualdi, *Introductory Combinatorics,* 3d ed., Prentice-Hall, Englewood Cliffs, NJ, 1997.

[BuHa90] F. Buckley and F. Harary, *Distance in Graphs,* Addison-Wesley, Redwood City, CA, 1990.

[Ca78] L. Carroll, *Symbolic Logic,* Crown, New York, 1978.

[ChGhMo79] V. Chachra, P. M. Ghare, and J. M. Moore, *Applications of Graph Theory Algorithms,* North-Holland, New York, 1979.

[Ch77] G. Chartrand, *Graphs as Mathematical Models,* Prindle, Weber & Schmidt, Boston, 1977.

[ChLe96] G. Chartrand and L. Lesniak, *Graphs and Digraphs,* 3d ed., Wadsworth, Belmont, CA, 1996.

[ChOe93] G. Chartrand and O. R. Oellermann, *Applied Algorithmic Graph Theory,* McGraw-Hill, New York, 1993.

[ClMe94] W. F. Clocksin and C. S. Mellish, *Programming in Prolog,* 4th ed., Springer-Verlag, New York, 1994.

[Co78] D. I. A. Cohen, *Basic Techniques of Combinatorial Theory,* Wiley, New York, 1978.

[Co00] D. Comer, Internetworking with TCP/IP, *Principles, Protocols, and Architecture,* vol. 1, 4th ed., Prentice-Hall, Englewood Cliffs, NJ, 2000.

[CoLeRiSt01] T. H. Cormen, C. E. Leierson, R. L. Rivest, and Stein, *Introduction to Algorithms,* 2d ed., MIT Press, Cambridge, MA, 2001.

[CrPo01] Richard Crandall and Carl Pomerance, *Prime Numbers: a Computational Perspective,* Springer-Verlag, New York, 2001.

[Cu01] Antonella Cupillari, *The Nuts and Bolts of Proofs,* 2d ed., Academic Press, San Diego, 2001.

[Cu84] C. W. Curtis, *Linear Algebra,* Springer-Verlag, New York, 1984.

[Da82] C. J. Date, *An Introduction to Database Systems,* 3d ed., Addison-Wesley, Reading, MA, 1982.

[De82] D. E. R. Denning, *Cryptography and Data Security,* Addison-Wesley, Reading, MA, 1982.

[DeDeQu81] P. J. Denning, J. B. Dennis, and J. E. Qualitz, *Machines, Languages, and Computation,* Prentice-Hall, Englewood Cliffs, NJ, 1981.

[De74] N. Deo, *Graph Theory with Applications to Engineering and Computer Science,* Prentice-Hall, Englewood Cliffs, NJ, 1974.

[De84] A. K. Dewdney, "Computer Recreations," *Scientific American,* 251, no. 2 (August 1984), 19–23; 252, no. 3 (March 1985), 14–23; 251, no. 4 (April 1985), 20–30.

[De93] A. K. Dewdney, *The New Turing Omnibus: Sixty-Six Excursions in Computer Science,* W. H. Freeman, New York, 1993.

[Di59] E. Dijkstra, "Two Problems in Connexion with Graphs," *Numerische Mathematik*, 1 (1959), 269–271.

[Eu53] L. Euler, "The Koenigsberg Bridges," *Scientific American*, 189, no. 1 (July 1953), 66–70.

[Ev73] S. Even, *Algorithmic Combinatorics*, Macmillan, New York, 1973.

[Ev79] S. Even, *Graph Algorithms*, Computer Science Press, Rockville, MD, 1979.

[Fe68] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1, 3d ed., Wiley, New York, 1968.

[Fo92] L. R. Foulds, *Graph Theory Applications*, Springer-Verlag, New York, 1992.

[GaJo79] Michael R. Garey and David S. Johnson, *Computers and Intractability: A Guide to NP-Completeness*, Freeman, New York, 1979.

[Gi85] A. Gibbons, *Algorithmic Graph Theory*, Cambridge University Press, Cambridge, England, 1985.

[Gi70] C. C. Gillispie, ed., *Dictionary of Scientific Biography*, Scribner's, New York, 1970.

[Go94] S. W. Golomb, *Polyominoes*, Princeton University Press, Princeton, NJ, 1994.

[Go84] G. H. Gonnet, *Handbook of Algorithms and Data Structures*, Addison-Wesley, London, 1984.

[GoHe77] S. E. Goodman and S. T. Hedetniemi, *Introduction to the Design and Analysis of Algorithms*, McGraw-Hill, New York, 1977.

[GoGo78] C. C. Gotlieb and L. R. Gotlieb, *Data Types and Structures*, Prentice-Hall, Englewood Cliffs, NJ, 1978.

[GrHe85] R. L. Graham and P. Hell, "On the History of the Minimum Spanning Tree Problem," *Annals of the History of Computing*, 7 (1985), 43–57.

[GrKnPa94] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, 2d ed., Addison-Wesley, Reading, MA, 1994.

[GrRoSp90] Ronald L. Graham, Bruce L. Rothschild, and Joel H. Spencer, *Ramsey Theory*, 2d ed., Wiley, New York, 1990.

[GrWa77] J. E. Graver and M. E. Watkins, *Combinatorics with Emphasis on the Theory of Graphs*, Springer-Verlag, New York, 1977.

[GrSc93] D. Gries and F. B. Schneider, *A Logical Approach to Discrete Math*, Springer-Verlag, New York, 1993.

[Gr94] R. P. Grimaldi, *Discrete and Combinatorial Mathematics*, 3d ed., Addison-Wesley, Reading, MA, 1994.

[GrYe99] J. L. Gross and J. Yellen, *Graph Theory and Its Applications*, CRC Press, Boca Raton, FL, 1999.

[Gr90] Jerrold W. Grossman, *Discrete Mathematics: An Introduction to Concepts, Methods, and Applications*, Macmillan, New York, 1990.

[Gr97] J. Gruska, *Foundations of Computing*, International Thomsen Computer Press, London, 1997.

[Ha60] P. R. Halmos, *Naive Set Theory*, D. Van Nostrand, New York, 1960.

[Ha80] R. W. Hamming, *Coding and Information Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[Ha69] F. Harary, *Graph Theory*, Addison-Wesley, Reading, MA, 1969.

[HaWr79] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 5th ed., Oxford University Press, Oxford, England, 1979.

[Ha87] D. Harel, *Algorithmics, The Spirit of Computing*, Addison-Wesley, Reading, MA, 1987.

[Ha00a] Brian Hayes, "Graph-Theory in Practice, Part I," *American Scientist*, 88, no. 1 (2000), 9–13.

[Ha00b] Brian Hayes, "Graph-Theory in Practice, Part II," *American Scientist*, 88, no. 2 (2000), 104–109.

[Ha93] John P. Hayes, *Introduction to Digital Logic Design*, Addison-Wesley, Reading, MA, 1993.

[He77] F. Hennie, *Introduction to Computability*, Addison-Wesley, Reading, MA, 1977.

[He88] R. Herken, *The Universal Turing Machine, A Half-Century Survey*, Oxford University Press, New York, 1988.

[Ho99] D. Hofstadter, *Gödel, Escher, Bach: An Internal Golden Braid*, Basic Books, New York, 1999.

[Ho66] F. E. Hohn, *Applied Boolean Algebra*, 2d ed., Macmillan, New York, 1966.

[HoUl79] J. E. Hopcroft and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley, Reading, MA, 1979.

[HoMo76] D. Hopkin and B. Moss, *Automata*, Elsevier, North-Holland, New York, 1976.

[HoSa82] E. Horowitz and S. Sahni, *Fundamentals of Computer Algorithms*, Computer Science Press, Rockville, MD, 1982.

[Hu82] T. C. Hu, *Combinatorial Algorithms*, Addison-Wesley, Reading, MA, 1982.

[In93] D. C. Ince, *An Introduction to Discrete Mathematics, Formal System Specification, and Z*, 2d ed., Oxford, New York, 1993.

[Ka53] M. Karnaugh, "The Map Method for Synthesis of Combinatorial Logic Circuits," *Transactions of the AIEE*, part I, 72 (1953), 593–599.

[Ka93] Randy H. Katz, *Contemporary Logic Design*, Addison-Wesley, Reading, MA, 1993.

[Kl56] S. C. Kleene, "Representation of Events by Nerve Nets," in *Automata Studies*, 3–42, Princeton University Press, Princeton, NJ, 1956.

[Kn77] D. E. Knuth, "Algorithms," *Scientific American*, 236, no. 4 (April 1977), 63–80.

[Kn97a] D. E. Knuth, *The Art of Computer Programming, Vol. I: Fundamental Algorithms*, 3d ed., Addison-Wesley, Reading, MA, 1997a.

[Kn97b] D. E. Knuth, *The Art of Computer Programming, Vol. II: Seminumerical Algorithms*, 3d ed., Addison-Wesley, Reading, MA, 1997b.

[Kn98] D. E. Knuth, *The Art of Computer Programming, Vol. III: Sorting and Searching,* 2d ed., Addison-Wesley, Reading, MA, 1998.

[Ko86] Z. Kohavi, *Switching and Finite Automata Theory,* 2d ed., McGraw-Hill, New York, 1986.

[KrSt98] Donald H. Kreher and Douglas R. Stinson, *Combinatorial Algorithms: Generation, Enumeration, and Search,* CRC Press. Boca Raton, FL 1998.

[Kr87] L. Kronsjö, *Algorithms: Their Complexity and Efficiency,* 2d ed., Wiley, New York, 1987.

[Kr56] J. B. Kruskal, "On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem," *Proceedings of the AMS,* 1 (1956), 48–50.

[Le64] D. H. Lehmer, "The Machine Tools of Combinatorics," in E. F. Beckenbach (ed.), *Applied Combinatorial Mathematics,* Wiley, New York, 1964.

[Le77] W. J. LeVeque, *Fundamentals of Number Theory,* Addison-Wesley, Reading, MA, 1977.

[LePa81] H. R. Lewis and C. H. Papadimitriou, *Elements of the Theory of Computation,* Prentice-Hall, Englewood Cliffs, NJ, 1981.

[LiLi81] Y. Lin and S. Y. T. Lin, *Set Theory with Applications,* 2d ed., Mariner, Tampa, FL, 1981.

[Li68] C. L. Liu, *Introduction to Combinatorial Mathematics,* McGraw-Hill, New York, 1968.

[Li85] C. L. Liu, *Elements of Discrete Mathematics,* 2d ed., McGraw-Hill, New York, 1985.

[Lo79] L. Lovász, *Combinatorial Problems and Exercises,* North-Holland, Amsterdam, 1979.

[Lu91] E. Lucas, *Récreations Mathématiques,* Gauthier-Villars, Paris, 1891.

[Ma91] J. C. Martin, *Introduction to Languages and the Theory of Computation,* McGraw-Hill, New York, 1991.

[Ma93] H. F. Mattson, Jr., *Discrete Mathematics with Applications,* Wiley, New York, 1993.

[Mc56] E. J. McCluskey, Jr., "Minimization of Boolean Functions," *Bell System Technical Journal,* 35 (1956), 1417–1444.

[Mc90] J. A. McHugh, *Algorithmic Graph Theory,* Prentice-Hall, Englewood Cliffs, NJ, 1990.

[Mc82] R. McNaughton, *Elementary Computability, Formal Languages, and Automata,* Prentice-Hall, Englewood Cliffs, NJ, 1982.

[Me55] G. H. Mealy, "A Method for Synthesizing Sequential Circuits," *Bell System Technical Journal,* 34 (1955), 1045–1079.

[Me64] E. Mendelson, *Introduction to Mathematical Logic,* Van Nostrand Reinhold, New York, 1964.

[MeOoVa97] A. J. Menezes, P. C. van Oorschot, S. A. Vanstone, *Handbook of Applied Cryptography,* CRC Press, Boca Raton, FL, 1997.

[MiRo91] J. G. Michaels and K. H. Rosen, *Applications of Discrete Mathematics,* McGraw-Hill, New York, 1991.

[Mo69] J. R. Monk, *Introduction to Set Theory,* McGraw-Hill, New York, 1969.

[Mo56] E. F. Moore, "Gedanken-Experiments on Sequential Machines," in *Automata Studies,* 129–153, Princeton University Press, Princeton, NJ, 1956.

[Na00] Paul J. Nabin, *Duelling Idiots and Other Probability Puzzlers,* Princeton University Press, Princeton, NJ, 2000.

[Ne85] C. V. Negoita, *Expert Systems and Fuzzy Systems,* Benjamin Cummings, Menlo Park, CA, 1985.

[NiMa95] Ulf Nilsson and Jan Maluszynski, *Logic, Programming, and Prolog,* 2d ed., Wiley, Chichester, England, 1995.

[Or63] O. Ore, *Graphs and Their Uses,* Mathematical Association of America, Washington, D.C., 1963.

[Pe87] A. Pelc, "Solution of Ulam's Problem on Searching with a Lie," *Journal of Combinatorial Theory,* Series A, 44 (1987), 129–140.

[PoSh81] I. Pohl and A. Shaw, *The Nature of Computation: An Introduction to Computer Science,* Computer Science Press, Rockville, MD, 1981.

[Po62] George Pólya, *Mathematical Discovery,* Vols. 1 and 2, Wiley, New York, 1962.

[Po71] George Pólya, *How to Solve It,* Princeton University Press, Princeton, NJ, 1971.

[Po90] George Pólya, *Mathematics and Plausible Reasoning,* Princeton University Press, Princeton, NJ, 1990.

[PoTaWo83] G. Pólya, R. E. Tarjan, and D. R. Woods, *Notes on Introductory Combinatorics,* Birkhäuser, Boston, 1983.

[Pr57] R. C. Prim, "Shortest Connection Networks and Some Generalizations," *Bell System Technical Journal,* 36 (1957), 1389–1401.

[PuBr85] P. W. Purdom, Jr. and C. A. Brown, *The Analysis of Algorithms,* Holt, Rinehart & Winston, New York, 1985.

[Qu52] W. V. Quine, "The Problem of Simplifying Truth Functions," *American Mathematical Monthly,* 59 (1952), 521–531.

[Qu55] W. V. Quine, "A Way to Simplify Truth Functions," *American Mathematical Monthly,* 62 (1955), 627–631.

[Ra62] T. Rado, "On Non-Computable Functions," *Bell System Technical Journal,* (May 1962), 877–884.

[Ra92] Gregory J. E. Rawlins, *Compared to What? An Introduction to the Analysis of Algorithms,* Computer Science Press, New York, 1992.

[ReNiDe77] E. M. Reingold, J. Nievergelt, and N. Deo, *Combinatorial Algorithms: Theory and Practice,* Prentice-Hall, Englewood Cliffs, NJ, 1977.

[Ri58] J. Riordan, *An Introduction to Combinatorial Analysis,* Wiley, New York, 1958.

[Ri68] J. Riordan, *Combinatorial Identities,* Wiley, New York, 1968.

[RiShAd78] R. Rivest, A. Shamir, and L. Adleman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems," *Communications of the Association for Computing Machinery,* 31, no. 2 (1978), 120–128.

[Ro86] E. S. Roberts, *Thinking Recursively,* Wiley, New York, 1986.

[Ro76] F. S. Roberts, *Discrete Mathematics Models,* Prentice-Hall, Englewood Cliffs, NJ, 1976.

[Ro84] F. S. Roberts, *Applied Combinatorics,* Prentice-Hall, Englewood Cliffs, NJ, 1984.

[RoFo80] D. F. Robinson and L. R. Foulds, *Digraphs: Theory and Techniques,* Gordon and Breach, New York, 1980.

[Ro84a] J. S. Rohl, *Recursion via Pascal,* Cambridge University Press, Cambridge, England, 1984.

[Ro00] K. H. Rosen, *Elementary Number Theory and Its Applications,* 4th ed., Addison-Wesley, Reading, MA, 2000.

[Ro00a] K. H. Rosen, *Handbook of Discrete and Combinatorial Mathematics,* CRC Press, Boca Raton, FL, 2000.

[Ro01] Sheldon M. Ross, *A First Course in Probability Theory,* 6th ed., Prentice-Hall, Englewood Cliffs, NJ, 2001.

[Ro02] Sheldon M. Ross, *Probability Models for Computer Science,* Harcourt/Academic Press, 2002.

[Ro59] B. Roy, "Transitivité et Connexité," *C.R. Acad. Sci. Paris,* 249 (1959), 216.

[Ry63] H. Ryser, *Combinatorial Mathematics,* Mathematical Association of America, Washington, D.C., 1963.

[SaKa86] T. L. Saaty and P. C. Kainen, *The Four-Color Problem: Assaults and Conquest,* Dover, New York, 1986.

[Sa85] S. Sahni, *Concepts in Discrete Mathematics,* Camelot, Minneapolis, 1985.

[SePi89] J. Seberry and J. Pieprzyk, *Cryptography: An Introduction to Computer Security,* Prentice-Hall, Englewood Cliffs, NJ, 1989.

[Se88] R. Sedgewick, *Algorithms,* 2d ed., Addison-Wesley, Reading, MA, 1988.

[SiHs88] H. J. Siegel and W. T. Hsu, "Interconnection Networks," in *Computer Architectures,* V. M. Milutinovic (ed.), North-Holland, New York, 1988, pp. 225–264.

[Si66] A. Sinkov, *Elementary Cryptanalysis,* Mathematical Association of America, Washington, DC, 1966.

[SlPl95] N. J. A. Sloane and S. Plouffe, *The Encyclopedia of Integer Sequences,* Academic Press, New York, 1995.

[Sm78] Raymond Smullyan, *What Is the Name of This Book?: The Riddle of Dracula and Other Logical Puzzles,* Prentice-Hall, Englewood Cliffs, NJ, 1978.

[Sm92] Raymond Smullyan, *Lady or the Tiger? And Other Logic Puzzles Including a Mathematical Novel That Features Godel's Great Discovery,* Times Book, New York, 1992.

[Sm98] Raymond Smullyan, *The Riddle of Scheherazade: And Other Amazing Puzzles, Ancient & Modern,* Harvest Book, Fort Washington, PA, 1998.

[So01] Daniel Solow, *How to Read and Do Proofs: An Introduction to Mathematical Thought Processes,* Wiley, New York, 2001.

[So61] I. S. Sominskii, *Method of Mathematical Induction,* Blaisdell, New York, 1961.

[Sp94] M. Spivak, *Calculus,* 3d ed., Publish or Perish, Wilmington, DE, 1994.

[StMc77] D. Stanat and D. F. McAllister, *Discrete Mathematics in Computer Science,* Prentice-Hall, Englewood Cliffs, NJ, 1977.

[St78] H. M. Stark, *An Introduction to Number Theory,* MIT Press, Cambridge, MA, 1978.

[St02] Douglas R. Stinson, *Cryptography, Theory and Practice,* 2d ed., Chapman and Hall/CRC, Boca Raton, FL, 2002.

[St94] P. K. Stockmeyer, "Variations on the Four-Post Tower of Hanoi Puzzle," *Congressus Numerantrum* 102 (1994), 3–12.

[St74] R. R. Stoll, *Sets, Logic, and Axiomatic Theories,* 2d ed., W. H. Freeman, San Francisco, 1974.

[St88] G. W. Strang, *Linear Algebra and Its Applications,* 3d ed., Harcourt Brace Jovanovich, San Diego, 1988.

[Su87] P. Suppes, *Introduction to Logic,* D. Van Nostrand, Princeton, NJ, 1987.

[Ta83] R. E. Tarjan, *Data Structures and Network Algorithms,* Society for Industrial and Applied Mathematics, Philadelphia, 1983.

[ThFi96] G. B. Thomas and R. L. Finney, *Calculus and Analytic Geometry,* 9th ed., Addison-Wesley, Reading, MA, 1996.

[TrMa75] J. P. Tremblay and R. P. Manohar, *Discrete Mathematical Structures with Applications to Computer Science,* McGraw-Hill, New York, 1975.

[Tu02] Alan Tucker, *Applied Combinatorics,* 4th ed., Wiley, New York, 2002.

[Ve52] E. W. Veitch, "A Chart Method for Simplifying Truth Functions," *Proceedings of the ACM* (1952), 127–133.

[Ve94] David Velleman, *How to Prove It: A Structured Approach,* Cambridge University Press, New York, 1994.

[Vi71] N. Y. Vilenkin, *Combinatorics,* Academic Press, New York, 1971.

[Wa80] M. Wand, *Induction, Recursion, and Programming,* North-Holland, New York, 1980.

[Wa62] S. Warshall, "A Theorem on Boolean Matrices," *Journal of the ACM,* 9 (1962), 11–12.

[Wi86] Herbert S. Wilf, *Algorithms and Complexity,* Prentice-Hall, Englewood Cliffs, NJ, 1986.

[Wi85] S. G. Williamson, *Combinatorics for Computer Science,* Computer Science Press, Rockville, MD, 1985.

[Wi85] R. J. Wilson, *Introduction to Graph Theory,* 3d ed., Longman, Essex, England, 1985.

[WiBe79] R. J. Wilson and L. W. Beineke, *Applications of Graph Theory,* Academic Press, London, 1979.

[WiWa90] R. J. Wilson and J. J. Watkins, *Graphs, An Introductory Approach,* Wiley, New York, 1990.

[Wi76] N. Wirth, *Algorithms + Data Structures = Programs,* Prentice-Hall, Englewood Cliffs, NJ, 1976.

[Wi84] N. Wirth, "Data Structures and Algorithms," *Scientific American,* 251 (September 1984), 60–69.

[Wo98] Robert S. Wolf, *Proof, Logic, and Conjecture: The Mathematician's Toolbox,* New York, 1998.

[Wo87] D. Wood, *Theory of Computation,* Harper & Row, New York, 1987.

[Zi91] H. J. Zimmermann, *Fuzzy Set Theory and Its Applications,* 2d ed., Kluwer, Boston, 1991.